```
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Tosser-Klopp G
            Genetique Animale
            Institut National de la Recherche Agronomique
            Chemin de Borde-Rouge - Auzeville BP27, 31326 Castanet-Tolosan
            cedex, FRANCE
            Tel: 33 (0) 5.61.28.51.14
            Fax: 33 (0) 5.61.28.53.08
            Email: tosser@toulouse.inra.fr
            Clone distribution: AGENAE Resource centre. Francois PIUMI,
            Francois.Piumi@jouy.inra.fr, INRA, CEA Radiobiologie et Etude du
            genome (LREG), Domaine de Vilvert, 78352, Jouy-en-Josas cedex,
            FRANCE, +33 (0) 1.34.65.28.02, +33 (0) 1.34.65.22.73
            Sequence cleaned of vector, adaptator and repetitions. Contact us
            at sigenasupport@jouy.inra.fr to obtain the chromatogram of this
            sequence.
            Plate: 0029  row: b  column: 13.
FEATURES             Location/Qualifiers
     source          1. .530
                     /organism="Sus scrofa"
                     /mol_type="mRNA"
                     /db_xref="taxon:9823"
                     /clone="scac0029i.b.13"
                     /tissue_type="mixed"
                     /clone_lib="Sus Scrofa library (scac)"
                     /note="Vector: pT7T3D-pac vector; tissues: adipose tissue,
                     brain, kidney, liver, muscle, ovary, testis,  heart,
                     hypothalamus, pancreas, skin, spleen, thymus, placenta,
                     pituitary  gland, seminal vesicle, small intestine,
                     uterus, adrenals, bulbo uretral  gland, cerebral trunk,
                     epididymis, female gonad, gall-bladder,  hippocampus,
                     large intestine, male gonad, melanocytes, stomach, udder"
ORIGIN

Alignment Scores:
Pred. No.:                4.92e-43        Length:          530
Score:                    66.00           Matches:         99
Percent Similarity:       98.02%          Conservative:    0
Best Local Similarity:    98.02%          Mismatches:      1
Query Match:              4.62%           Indels:          2
DB:                       13              Gaps:            0

US-09-787-097-12 (1-1429) x BX669805 (1-530)

Qy        146 ThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeuIleGluGlyGln 165
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        429 ACAGATGGACCTGGAAATTATAAATATAAAACGAAGTGCACATGGCTCATTGAAGGACAG 370

Qy        166 ProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHis 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 CCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCAT 310

Qy        186 LeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaAlaPheSerGlyLeu 205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 TTATATGTTTATGATGGGGACTCAATTTATGCACCTCTAGTTGCTGCCTTTAGCGGCCTT 250

Qy        206 IleValProGluArgAspGly-AsnGluThrValProGluValValAlaThrSerGlyTy 225
              |||||||||||||||||||    ||||||||||||||||||||||||||||||||||||
Db        249 ATCGTTCCTGAGAGAGACAG-CAACGAGACTGTCCCAGAGGTCGTTGCCACATCAGGTTA 191

Qy        225 rAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThrTy 245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 TGCCCTGCTGCATTTTTTTAGCGATGCTGCTTATAATTTGACTGGATTTAATATCACTTA 131

Qy        245 r 245
              |
Db        130 C 130

RESULT 75
BU506373
LOCUS       BU506373                964 bp    mRNA    linear   EST 12-SEP-2002
DEFINITION  AGENCOURT_10015485 NIH_MGC_94 Mus musculus cDNA clone IMAGE:6494441
            5', mRNA sequence.
ACCESSION   BU506373
VERSION     BU506373.1  GI:22812606
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 964)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: The Cepko Laboratory
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM14050  row: k  column: 18
            High quality sequence stop: 592.
FEATURES             Location/Qualifiers
     source          1. .964
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:6494441"
                     /tissue_type="retina"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_94"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 3.3 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:                5.7e-42         Length:          964
Score:                    65.00           Matches:         78
Percent Similarity:       98.73%          Conservative:    0
Best Local Similarity:    98.73%          Mismatches:      0
Query Match:              4.55%           Indels:          1
DB:                       13              Gaps:            0

US-09-787-097-12 (1-1429) x BU506373 (1-964)

Qy        842 MetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLys 861
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        529 ATGCAATCATCTCAAAGTATGTCCAAGCTCACTCTGACTCCATGGGTTGGTCTTCGGAAG 588

Qy        862 IleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGln 881
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        589 ATCAATGTGTCTTACTGGTGCTGGGAGGATATGTCTCCATTCACAAATAGTTTGCTGCAG 648

Qy        882 TrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThr 901
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        649 TGGATGCCATCTGAGCCCAGTGATGCTGGCTTCTGTGGGATCTTGTCAGAGCCTAGTACT 708

Qy        902 Arg-GlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGlu 919
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        709 CGGGGGATTAAAGGCTGCAACCTGCATCAACCCTCTCAATGGCAGCGTCTGTGAA 763


Search completed: March  1, 2004, 12:54:55
Job time : 6985 secs
```

This Page Blank (uspto)

> O <
O| |O IntelliGenetics
> O <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file us-09-787-097-12.res made by jdelaval on Sat 17 Jul 104 6:56:22-PST.

Query sequence being compared:US-09-787-097-12 (1-1429)
Number of sequences searched: 4
Number of scores above cutoff: 3

Results of the initial comparison of US-09-787-097-12 (1-1429) with:
File : seqs-2-10-18.pep



PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	32
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	1248	1268	88.05

Times:	CPU	Total Elapsed
	00:00:00.00	00:00:00.00

Number of residues: 3825
Number of sequences searched: 4
Number of scores above cutoff: 3

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
1. US-09-787-097-10	Sequence 10, Application	1355	1325	1331	0.87	0
2. US-09-787-097-18	Sequence 18, Application	1272	1267	1268	0.22	0
	**** 1 standard deviation below mean ****					
3. US-09-787-097-2	Sequence 2, Application US	1198	1152	1159	-1.09	0

1. US-09-787-097-12 (1-1429)
US-09-787-097-10 Sequence 10, Application US/09787097

Sequence 10, Application US/09787097
GENERAL INFORMATION:
APPLICANT: Duke-Cohan, Jonathan S.
APPLICANT: Schlossman, Stuart F.
TITLE OF INVENTION: REGULATION OF IMMUNE RESPONSES BY ATTRACTIN
FILE REFERENCE: 00530-089002
CURRENT APPLICATION NUMBER: US/09/787,097
CURRENT FILING DATE: 2001-03-13
PRIOR APPLICATION NUMBER: PCT/US99/20948
PRIOR FILING DATE: 1999-09-14
PRIOR APPLICATION NUMBER: US 60/100,137
PRIOR FILING DATE: 1998-09-14
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 10
LENGTH: 1355
TYPE: PRT
ORGANISM: Homo sapiens

Initial Score	=	1325	Optimized Score	=	1331	Significance	= 0.87
Residue Identity	=	98%	Matches	=	1332	Mismatches	= 19
Gaps	=	4	Conservative Substitutions				= 0

```
          10         20         30         40         50         60         70
MVAAAAATEARLRRRTAATAALAGRSGGPHWDWDVTRAGRPGLGAGLRLPRLLSPPLRPRLLLLLLLLPPPL

          80         90        100        110        120        130        140
LLLLLPCEAEAAAAAAAAVSGSAAAEAKECDR----PCVNGGRCNPGTGQCVCPAGWVGEQCQHCGGRFRLTG
          | |||  |       ||  |     |     ||||||||||||||||||||||||||||||||||||||
          MVAAAAATEARLRRRTAATAALAGRSGGPHCVNGGRCNPGTGQCVCPAGWVGEQCQHCGGRFRLTG
          X        10         20         30         40         50         60

         150        160        170        180        190        200        210
SSGFVTDGPGNYKYKTKCTWLIEGQPNRIMRLRFNHFATECSWDHLYVYDGDSIYAPLVAAFSGLIVPERDG
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
SSGFVTDGPGNYKYKTKCTWLIEGQPNRIMRLRFNHFATECSWDHLYVYDGDSIYAPLVAAFSGLIVPERDG
  70         80         90        100        110        120        130

     220        230        240        250        260        270        280
NETVPEVVATSGYALLHFFSDAAYNLTGFNITYSFDMCPNNCSGRGECKISNSSETVECECSENWKGEACDI
|||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
NETVPEVVATSGYALLHFFSDAAYNLTGFNITYSFDMCPNNCSGRGECKISNSSDTVECECSENWKGEACDI
140        150        160        170        180        190        200        210

   290        300        310        320        330        340        350
PHCTDNCGFPHRGICNSSDVRGCSCFSDWQGPGCSVPVPANQSFWTREEYSNLKLPRASHKAVVNGNIMWVV
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
PHCTDNCGFPHRGICNSSDVRGCSCFSDWQGPGCSVPVPANQSFWTREEYSNLKLPRASHKAVVNGNIMWVV
         220        230        240        250        260        270        280

 360        370        380        390        400        410        420
GGYMFNHSDYNMVLAYDLASREWLPLNRSVNNVVVRYGHSLALYKDKIYMYGGKIDPTGNVTNELRVFHIHN
|||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
GGYMFNHSDYNMVLAYDLASREWLPLNRSVNNVVVRYGHSLALYKDKIYMYGGKIDSTGNVTNELRVFHIHN
      290        300        310        320        330        340        350

430        440        450        460        470        480        490        500
ESWVLLTPKAKEQYAVVGHSAHIVTLKNGRVVMLVIFGHCPLYGYISNVQEYDLDKNTWSILHTQGALVQGG
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ESWVLLTPKAKEQYAVVGHSAHIVTLKNGRVVMLVIFGHCPLYGYISNVQEYDLDKNTWSILHTQGALVQGG
   360        370        380        390        400        410        420
```

```
        510       520       530       540       550       560       570
YGHSSVYDHRTRALYVHGGYKAFSANKYRLADDLYRYDVDTQMWTILKDSRFFRYLHTAVIVSGTMLVFGGN
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
YGHSSVYDHRTRALYVHGGYKAFSANKYRLADDLYRYDVDTQMWTILKDSRFFRYLHTAVIVSGTMLVFGGN
 430       440       450       460       470       480       490

      580       590       600       610       620       630       640
THNDTSMSHGAKCFSSDFMAYDIACDRWSVLPRPDLHHDVNRFGHSAVLHNSTMYVFGGFNSLLLSDILVFT
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
THNDTSMSHGAKCFSSDFMAYDIACDRWSVLPRPDLHHDVNRFGHSAVLHNSTMYVFGGFNSLLLSDILVFT
500       510       520       530       540       550       560       570

    650       660       670       680       690       700       710
SEQCDAHRSEAACLAAGPGIRCVWNTGSSQCISWALATDEQEEKLKSECFSKRTLDHDRCDQHTDCYSCTAN
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
SEQCDAHRSEAACLAAGPGIRCVWNTGSSQCISWALATDEQEEKLKSECFSKRTLDHDRCDQHTDCYSCTAN
        580       590       600       610       620       630       640

  720       730       740       750       760       770       780
TNDCHWCNDHCVPRNHSCSEGQISIFRYENCPKDNPMYYCNKKTSCRSCALDQNCQWEPRNQECIALPENIC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TNDCHWCNDHCVPRNHSCSEGQISIFRYENCPKDNPMYYCNKKTSCRSCALDQNCQWEPRNQECIALPENIC
      650       660       670       680       690       700       710

790       800       810       820       830       840       850       860
GIGWHLVGNSCLKITTAKENYDNAKLFCRNHNALLASLTTQKKVEFVLKQLRIMQSSQSMSKLTLTPWVGLR
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GIGWHLVGNSCLKITTAKENYDNAKLFCRNHNALLASLTTQKKVEFVLKQLRIMQSSQSMSKLTLTPWVGLR
   720       730       740       750       760       770       780

        870       880       890       900       910       920       930
KINVSYWCWEDMSPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLNGSVCERPANHSAKQCRTP
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
KINVSYWCWEDMSPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLNGSVCERPANHSAKQCRTP
 790       800       810       820       830       840       850

      940       950       960       970       980       990      1000
CALRTACGDCTSGSSECMWCSNMKQCVDSNAYVASPPFGQCMEWYTMSTCPPENCSGYCTCSHCLEQPGCGW
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CALRTACGDCTSGSSECMWCSNMKQCVDSNAYVASPPFGQCMEWYTMSTCPPENCSGYCTCSHCLEQPGCGW
860       870       880       890       900       910       920       930

   1010      1020      1030      1040      1050      1060      1070
CTDPSNTGKGKCIEGSYKGPVKMPSQAPTGNFYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSI
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTDPSNTGKGKCIEGSYKGPVKMPSQAPTGNFYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSI
        940       950       960       970       980       990      1000

 1080      1090      1100      1110      1120      1130      1140
CEKCENLTTGKHCETCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGDECQLCEVENRYQG
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CEKCENLTTGKHCETCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGDECQLCEVENRYQG
     1010      1020      1030      1040      1050      1060      1070

1150      1160      1170      1180      1190      1200      1210      1220
  NPLRGTCYYTLLIDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASKNFNLNITWAASFSAGTQAG
  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  NPLRGTCYYTLLIDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASKNFNLNITWAASFSAGTQAG
    1080      1090      1100      1110      1120      1130      1140

       1230      1240      1250      1260      1270      1280      1290
  EEMPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTWPIKIQIAFSQHSNFMDLVQFFVTFFSCFLS
  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  EEMPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTWPIKIQIAFSQHSNFMDLVQFFVTFFSCFLS
  1150      1160      1170      1180      1190      1200      1210

     1300      1310      1320      1330      1340      1350      1360
  LLLVAAVVWKIKQSCWASRRREQLLREMQQMASRPFASVNVALETDEEPPDLIGGSIKTVPKPIALEPCFGN
  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  LLLVAAVVWKIKQSCWASRRREQLLREMQQMASRPFASVNVALETDEEPPDLIGGSIKTVPKPIALEPCFGN
1220      1230      1240      1250      1260      1270      1280      1290
```

```
      1370      1380      1390      1400      1410      1420        X
KAAVLSVFVRLPRGLGGIPPPGQSGLAVASALVDISQQMPIVYKEKSGAVRNRKQQPPAQPGTCI
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
KAAVLSVFVRLPRGLGGIPPPGQSGLAVASALVDISQQMPIVYKEKSGAVRNRKQQPPAQPGTCI
      1300      1310      1320      1330      1340      1350    X
```

2. US-09-787-097-12 (1-1429)
   US-09-787-097-18 Sequence 18, Application US/09787097

Sequence 18, Application US/09787097
GENERAL INFORMATION:
 APPLICANT: Duke-Cohan, Jonathan S.
 APPLICANT: Schlossman, Stuart F.
 TITLE OF INVENTION: REGULATION OF IMMUNE RESPONSES BY ATTRACTIN
 FILE REFERENCE: 00530-089002
 CURRENT APPLICATION NUMBER: US/09/787,097
 CURRENT FILING DATE: 2001-03-13
 PRIOR APPLICATION NUMBER: PCT/US99/20948
 PRIOR FILING DATE: 1999-09-14
 PRIOR APPLICATION NUMBER: US 60/100,137
 PRIOR FILING DATE: 1998-09-14
 NUMBER OF SEQ ID NOS: 25
 SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 18
 LENGTH: 1272
 TYPE: PRT
 ORGANISM: Homo sapiens

Initial Score	= 1267	Optimized Score	= 1268	Significance	= 0.22
Residue Identity	= 99%	Matches	= 1268	Mismatches	= 4
Gaps	= 0	Conservative Substitutions			= 0

```
    X       10        20        30        40        50        60        70
    MVAAAAATEARLRRRTAATAALAGRSGGPHWDWDVTRAGRPGLGAGLRLPRLLSPPLRPRLLLLLLLLPPPL
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    MVAAAAATEARLRRRTAATAALAGRSGGPHWDWDVTRAGRPGLGAGLRLPRLLSPPLRPRLLLLLLLLPPPL
    X       10        20        30        40        50        60        70

            80        90       100       110       120       130       140
    LLLLLPCEAEAAAAAAAAVSGSAAAEAKECDRPCVNGGRCNPGTGQCVCPAGWVGEQCQHCGGRFRLTGSSGF
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    LLLLLPCEAEAAAAAAAAVSGSAAAEAKECDRPCVNGGRCNPGTGQCVCPAGWVGEQCQHCGGRFRLTGSSGF
            80        90       100       110       120       130       140

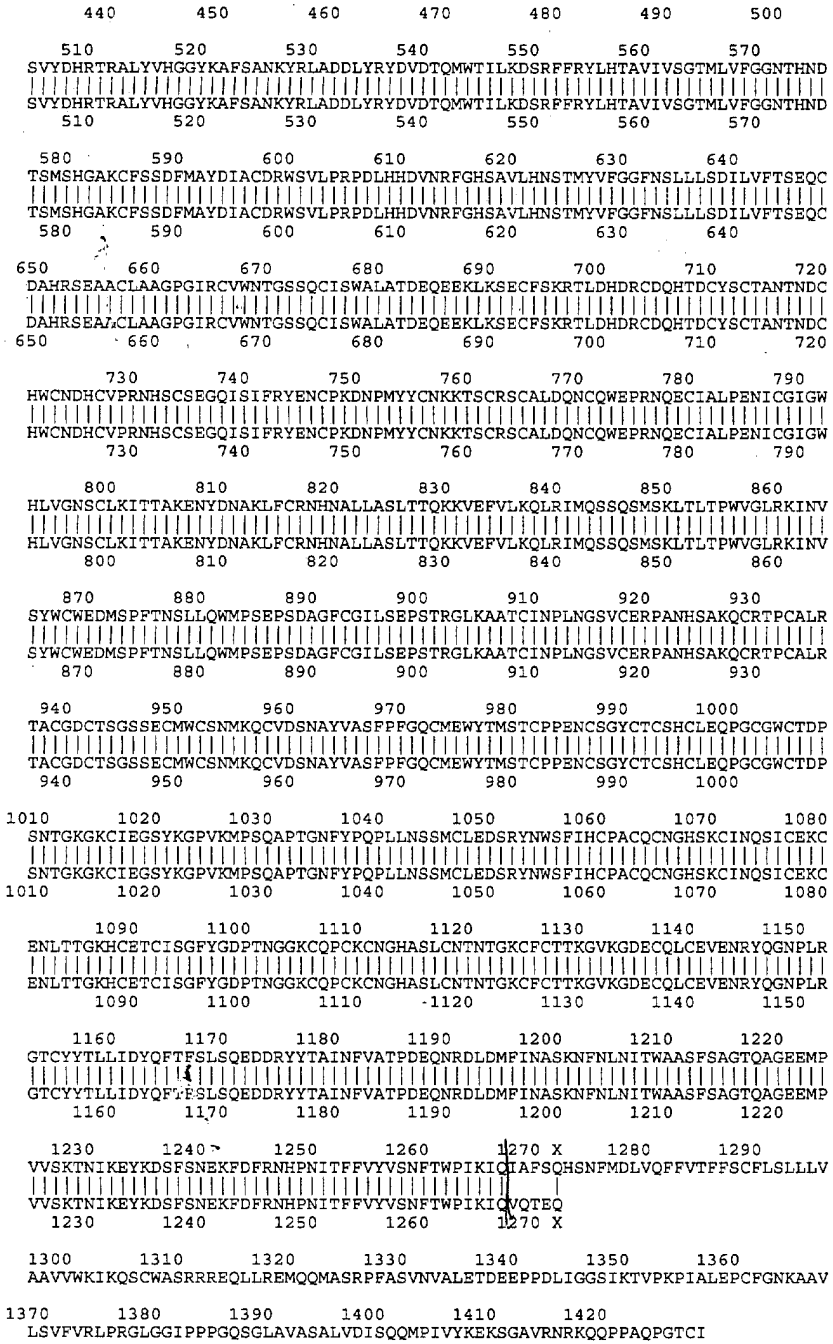
       150       160       170       180       190       200       210
    VTDGPGNYKYKTKCTWLIEGQPNRIMRLRFNHFATECSWDHLYVYDGDSIYAPLVAAFSGLIVPERDGNETV
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    VTDGPGNYKYKTKCTWLIEGQPNRIMRLRFNHFATECSWDHLYVYDGDSIYAPLVAAFSGLIVPERDGNETV
       150       160       170       180       190       200       210

     220       230       240       250       260       270       280
    PEVVATSGYALLHFFSDAAYNLTGFNITYSFDMCPNNCSGRGECKISNSSETVECECSENWKGEACDIPHCT
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    PEVVATSGYALLHFFSDAAYNLTGFNITYSFDMCPNNCSGRGECKISNSSETVECECSENWKGEACDIPHCT
     220       230       240       250       260       270       280

290       300       310       320       330       340       350       360
 DNCGFPHRGICNSSDVRGCSCFSDWQGPGCSVPVPANQSFWTREEYSNLKLPRASHKAVVNGNIMWVVGGYM
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DNCGFPHRGICNSSDVRGCSCFSDWQGPGCSVPVPANQSFWTREEYSNLKLPRASHKAVVNGNIMWVVGGYM
290       300       310       320       330       340       350       360

        370       380       390       400       410       420       430
 FNHSDYNMVLAYDLASREWLPLNRSVNNVVVRYGHSLALYKDKIYMYGGKIDPTGNVTNELRVFHIHNESWV
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 FNHSDYNMVLAYDLASREWLPLNRSVNNVVVRYGHSLALYKDKIYMYGGKIDPTGNVTNELRVFHIHNESWV
        370       380       390       400       410       420       430

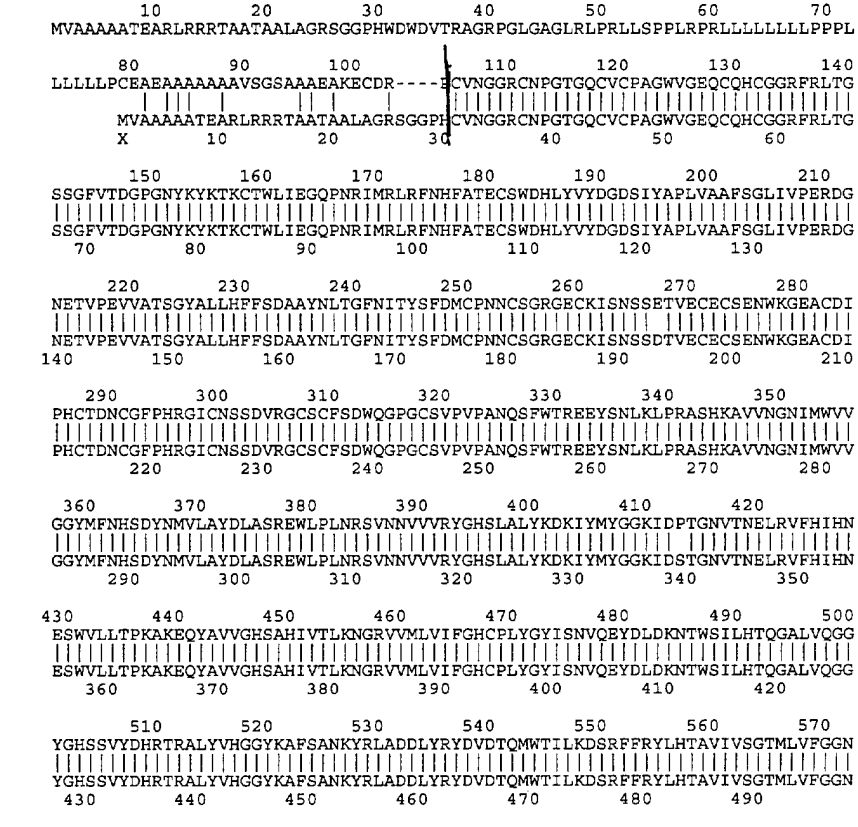
      440       450       460       470       480       490       500
 LLTPKAKEQYAVVGHSAHIVTLKNGRVVMLVIFGHCPLYGYISNVQEYDLDKNTWSILHTQGALVQGGYGHS
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 LLTPKAKEQYAVVGHSAHIVTLKNGRVVMLVIFGHCPLYGYISNVQEYDLDKNTWSILHTQGALVQGGYGHS
```



3. US-09-787-097-12 (1-1429)
US-09-787-097-2 Sequence 2, Application US/09787097

Sequence 2, Application US/09787097
GENERAL INFORMATION:
APPLICANT: Duke-Cohan, Jonathan S.
APPLICANT: Schlossman, Stuart F.
TITLE OF INVENTION: REGULATION OF IMMUNE RESPONSES BY ATTRACTIN
FILE REFERENCE: 00530-089002
CURRENT APPLICATION NUMBER: US/09/787,097
CURRENT FILING DATE: 2001-03-13
PRIOR APPLICATION NUMBER: PCT/US99/20948
PRIOR FILING DATE: 1999-09-14
PRIOR APPLICATION NUMBER: US 60/100,137
PRIOR FILING DATE: 1998-09-14
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 2
LENGTH: 1198
TYPE: PRT
ORGANISM: Homo sapiens

Initial Score = 1152 Optimized Score = 1159 Significance = -1.09
Residue Identity = 96% Matches = 1160 Mismatches = 34
Gaps = 4 Conservative Substitutions = 0



```
      580       590       600       610       620       630       640
 THNDTSMSHGAKCFSSDFMAYDIACDRWSVLPRPDLHHDVNRFGHSAVLHNSTMYVFGGFNSLLLSDILVFT
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 THNDTSMSHGAKCFSSDFMAYDIACDRWSVLPRPDSTMMSTDLAIPAVLHNSTMYVFGGFNSLLLSDILVFT
500       510       520       530       540       550       560       570

    650       660       670       680       690       700       710
 SEQCDAHRSEAACLAAGPGIRCVWNTGSSQCISWALATDEQEEKLKSECFSKRTLDHDRCDQHTDCYSCTAN
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 SEQCDAHRSEAACLAAGPGIRCVWNTGSSQCISWALATDEQEEKLKSECFSKRTLDHDRCDQHTDCYSCTAN
        580       590       600       610       620       630       640

  720       730       740       750       760       770       780
 TNDCHWCNDHCVPRNHSCSEGQISIFRYENCPKDNPMYYCNKKTSCRSCALDQNCQWEPRNQECIALPENIC
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 TNDCHWCNDHCVPRNHSCSEGQISIFRYENCPKDNPMYYCNKKTSCRSCALDQNCQWEPRNQECIALPENIC
      650       660       670       680       690       700       710

790       800       810       820       830       840       850       860
 GIGWHLVGNSCLKITTAKENYDNAKLFCRNHNALLASLTTQKKVEFVLKQLRIMQSSQSMSKLTLTPWVGLR
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 GIGWHLVGNSCLKITTAKENYDNAKLFCRNHNALLASLTTQKKVEFVLKQLRIMQSSQSMSKLTLTPWVGLR
    720       730       740       750       760       770       780

        870       880       890       900       910       920       930
 KINVSYWCWEDMSPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLNGSVCERPANHSAKQCRTP
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 KINVSYWCWEDMSPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLNGSVCERPANHSAKQCRTP
  790       800       810       820       830       840       850

      940       950       960       970       980       990      1000
 CALRTACGDCTSGSSECMWCSNMKQCVDSNAYVASFPFGQCMEWYTMSTCPPENCSGYCTCSHCLEQPGCGW
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 CALRTACGDCTSGSSECMWCSNMKQCVDSNAYVASFPFGQCMEWYTMSTCPPENCSGYCTCSHCLEQPGCGW
860       870       880       890       900       910       920       930

   1010      1020      1030      1040      1050      1060      1070
 CTDPSNTGKGKCIEGSYKGPVKMPSQAPTGNFYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSI
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 CTDPSNTGKGKCIEGSYKGPVKMPSQAPTGNFYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSI
       940       950       960       970       980       990      1000

 1080      1090      1100      1110      1120      1130      1140
 CEKCENLTTGKHCETCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGDECQLCEVENRYQG
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 CEKCENLTTGKHCETCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGDECQLCEVENRYQG
     1010      1020      1030      1040      1050      1060      1070

1150      1160      1170      1180      1190      1200      1210      1220
 NPLRGTCYYTLLIDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASKNFNLNITWAASFSAGTQAG
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 NPLRGTCYYTLLIDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASKNFNLNITWAASFSAGTQAG
   1080      1090      1100      1110      1120      1130      1140

        1230      1240      1250      1260      1270 X    1280      1290
 EEMPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTWPIKIQIAFSQHSNFMDLVQFFVTFFSCFLS
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 EEMPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTWPIKIQVQTEQ
 1150      1160      1170      1180      1190      X

     1300      1310      1320      1330      1340      1350      1360
 LLLVAAVVWKIKQSCWASRRREQLLREMQQMASRPFASVNVALETDEEPPDLIGGSIKTVPKPIALEPCFGN

   1370      1380      1390      1400      1410      1420
 KAAVLSVFVRLPRGLGGIPPPGQSGLAVASALVDISQQMPIVYKEKSGAVRNRKQQPPAQPGTCI
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2004, 16:07:54 ; Search time 9050.88 Seconds
(without alignments)
17225.392 Million cell updates/sec

Title: US-09-787-097-1
Perfect score: 3597
Sequence: 1 atggtggccgcagcggcggc.....aggtgcaaactgaacaatga 3597

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3597	100.0	3597	9	AF034957	AF034957 Homo sapi
2	3494.4	97.1	4072	6	AR164817	AR164817 Sequence
3	3487	96.9	3819	9	AF106861	AF106861 Homo sapi
4	3476.8	96.7	8589	6	AR164816	AR164816 Sequence
5	3100.8	86.2	4493	4	AF531101	AF531101 Bos tauru
6	2948.8	82.0	6083	10	AF116897	AF116897 Mus muscu
7	2947.8	82.0	4284	10	AB062913	AB062913 Mesocrice
8	2913.6	81.0	4313	10	AF119821	AF119821 Mus muscu
9	2909	80.9	8739	10	AB038387	AB038387 Rattus no
10	2907.2	80.8	4500	10	AB038388	AB038388 Rattus no
11	2845.2	79.1	8827	6	AR164807	AR164807 Sequence
12	2255	62.7	5683	10	AB093245	AB093245 Mus muscu
13	2028.8	56.4	2625	6	AR164818	AR164818 Sequence
14	1428.8	39.7	2419	6	AR164813	AR164813 Sequence
15	1364.4	37.9	6370	6	AR164815	AR164815 Sequence
16	1208.4	33.6	2578	9	AK000356	AK000356 Homo sapi
17	1203.2	33.5	4140	9	AY442317	AY442317 Homo sapi
18	1203.2	33.5	4463	9	AK127277	AK127277 Homo sapi
19	1129.4	31.4	3609	6	AX503776	AX503776 Sequence
20	879.8	24.5	7302	9	AB011106	AB011106 Homo sapi
21	872.4	24.3	5632	9	AB011120	AB011120 Homo sapi
22	829.2	23.1	5002	10	BC027764	BC027764 Mus muscu
23	817.8	22.7	1012	6	E26747	E26747 Structure a
24	543	15.1	543	6	AX778545	AX778545 Sequence
25	463	12.9	1051	6	AR164814	AR164814 Sequence
26	435.4	12.1	4997	10	BC050020	BC050020 Mus muscu
27	418.4	11.6	1818	9	BC047716	BC047716 Homo sapi
28	340	9.5	3297	10	BC030872	BC030872 Mus muscu
29	253.8	7.1	982	6	AX364943	AX364943 Sequence
30	250.8	7.0	2018	9	H009384S08	AF218896 Homo sapi
31	250.8	7.0	104907	9	HSDJ741H3	AL132773 Human DNA
c 32	250.8	7.0	155869	2	AC023420	AC023420 Homo sapi
c 33	250.8	7.0	167357	2	AC055771	AC055771 Homo sapi
34	250.8	7.0	206647	9	AP002898	AP002898 Homo sapi
c 35	246	6.8	246	6	AR273223	AR273223 Sequence
c 36	246	6.8	246	6	AR276804	AR276804 Sequence
c 37	246	6.8	246	6	AR407079	AR407079 Sequence
c 38	246	6.8	246	6	AX368256	AX368256 Sequence
39	234.4	6.5	3257	9	H009384S15	AF218903 Homo sapi
c 40	234.4	6.5	118540	2	AC013324	AC013324 Homo sapi
41	234.4	6.5	207433	6	AX326809	AX326809 Sequence
42	224.6	6.2	3589	9	H009384S11	AF218899 Homo sapi
43	222.6	6.2	1728	9	H009384S14	AF218902 Homo sapi
44	214.6	6.0	237	4	AF273610	AF273610 Sus scrof
45	214	5.9	1232	9	H009384S05	AF218893 Homo sapi

ALIGNMENTS

RESULT 1
AF034957
LOCUS AF034957 3597 bp mRNA linear PRI 30-SEP-1998
DEFINITION Homo sapiens secreted T cell activation protein Attractin (attractin) mRNA, complete cds.
ACCESSION AF034957
VERSION AF034957.1 GI:3676346
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3597)
AUTHORS Duke-Cohan,J.S., Gu,J., McLaughlin,D.F., Xu,Y., Freeman,G.J. and Schlossman,S.F.

```
  TITLE     Attractin (DPPT-L), a member of the CUB family of cell adhesion and
            guidance proteins, is secreted by activated human T lymphocytes and
            modulates immune cell interactions
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 95 (19), 11336-11341 (1998)
  MEDLINE   98409658
   PUBMED   9736737
REFERENCE   2  (bases 1 to 3597)
  AUTHORS   Duke-Cohan,J.S., Gu,J., Ao,Z., McLaughlin,D.F., Freeman,G.J. and
            Schlossman,S.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-NOV-1997) Division of Cancer, Immunology and AIDS,
            Dana Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
            USA
FEATURES             Location/Qualifiers
     source          1. .3597
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     gene            1. .3597
                     /gene="attractin"
     CDS             1. .3597
                     /gene="attractin"
                     /note="secreted T cell activation protein; Method:
                     conceptual translation with partial peptide sequencing;
                     DPPT-L"
                     /codon_start=1
                     /product="Attractin"
                     /protein_id="AAC61902.1"
                     /db_xref="GI:3676347"
                     /translation="MVAAAAATEARLRRRTAATAALAGRSGGPHCVNGGRCNPGTGQC
                     VCPAGWVGEQCQHCGGRFRLTGSSGFVTDGPGNYKYKTKCTWLIEGQPNRIMRLRFNH
                     FATECSWDHLYVYDGDSIYAPLVAAFSGLIVPERDGNETVPEVVATSGYALLHFFSDA
                     AYNLTGFNITYSFDMCPNNCSGRGECKISNSSDTVECECSENWKGEACDIPHCTDNCG
                     FPHRGICNSSDVRGCSCFSDWQGPGCSVPVPANQSFWTREEYSNLKLPRASHKAVVNG
                     NIMWVVGGYMFNHSDYNMVLAYDLASREWLPLNRSVNNVVVRYGHSLALYKDKIYMYG
                     GKIDSTGNVTNELRVFHIHNESWVLLTPKAKEQYAVVGHSAHIVTLKNGRVVMLVIFG
                     HCPLYGYISNVQEYDLDKNTWSILHTQGALVQGGYGHSSVYDHRTRALYVHGGYKAFS
                     ANKYRLADDLYRYDVDTQMWTILKDSRFFRYLHTAVIVSGTMLVFGGNTHNDTSMSHG
                     AKCFSSDFMAYDIACDRWSVLPRPDSTMMSTDLAIPAVLHNSTMYVFGGFNSLLLSDI
                     LVFTSEQCDAHRSEAACLAAGPGIRCVWNTGSSQCISWALATDEQEEKLKSECFSKRT
                     LDHDRCDQHTDCYSCTANTNDCHWCNDHCVPRNHSCSEGQISIFRYENCPKDNPMYYC
                     NKKTSCRSCALDQNCQWEPRNQECIALPENICGIGWHLVGNSCLKITTAKENYDNAKL
                     FCRNHNALLASLTTQKKVEFVLKQLRIMQSSQSMSKLTLTPWVGLRKINVSYWCWEDM
                     SPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLNGSVCERPANHSAKQCR
                     TPCALRTACGDCTSGSSECMWCSNMKQCVDSNAYVASFPFGQCMEWYTMSTCPPENCS
                     GYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKMPSQAPTGNFYPQPLLNSSM
                     CLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCENLTTGKHCETCISGFYGDPTNGG
                     KCQPCKCNGHASLCNTNTGKCFCTTKGVKGDECQLCEVENRYQGNPLRGTCYYTLLID
                     YQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASKNFNLNITWAASFSAGTQAG
                     EEMPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTWPIKIQVQTEQ"
ORIGIN

  Query Match           100.0%;  Score 3597;  DB 9;  Length 3597;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 3597;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 60

Qy         61 GCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGC 120

Qy        121 ACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGC 180

Qy        181 CGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATAC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATAC 240

Qy        241 AAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTC 300

Qy        301 AATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATT 360

Qy        361 TATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAG 420

Qy        421 ACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCT 480

Qy        481 GCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGC 540

Qy        541 TCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCT 600

Qy        601 GAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCT 660

Qy        661 CATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAG 720

Qy        721 GGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATAT 780

Qy        781 TCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGG 840

Qy        841 GTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTT 900

Qy        901 GCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGT 960

Qy        961 CATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACT 1020

Qy       1021 GGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTG 1080

Qy       1081 ACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTG 1140

Qy       1141 AAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATA 1200

Qy       1201 AGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGT 1260

Qy       1261 GCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTA 1320

Qy       1321 TACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      1321 TACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTC 1380
Qy      1381 TACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGT 1440
Db      1381 TACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGT 1440
Qy      1441 TACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACAC 1500
Db      1441 TACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACAC 1500
Qy      1501 AATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGAC 1560
Db      1501 AATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGAC 1560
Qy      1561 ATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATTCCACCATGATGTCAACAGAT 1620
Db      1561 ATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATTCCACCATGATGTCAACAGAT 1620
Qy      1621 TTGGCCATTCCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGT 1680
Db      1621 TTGGCCATTCCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGT 1680
Qy      1681 CTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAA 1740
Db      1681 CTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAA 1740
Qy      1741 GCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAG 1800
Db      1741 GCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAG 1800
Qy      1801 TGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTT 1860
Db      1801 TGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTT 1860
Qy      1861 TCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACA 1920
Db      1861 TCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACA 1920
Qy      1921 GCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGC 1980
Db      1921 GCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGC 1980
Qy      1981 TCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTAC 2040
Db      1981 TCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTAC 2040
Qy      2041 TACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAG 2100
Db      2041 TACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAG 2100
Qy      2101 CCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTG 2160
Db      2101 CCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTG 2160
Qy      2161 GTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTG 2220
Db      2161 GTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTG 2220
Qy      2221 TTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTT 2280
Db      2221 TTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTT 2280
Qy      2281 GTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACC 2340
Db      2281 GTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACC 2340
Qy      2341 CCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCA 2400
Db      2341 CCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCA 2400
Qy      2401 TTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGA 2460
Db      2401 TTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGA 2460
```

```
Qy      2461 ATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAAT 2520
Db      2461 ATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAAT 2520
Qy      2521 GGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCC 2580
Db      2521 GGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCC 2580
Qy      2581 TTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAAC 2640
Db      2581 TTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAAC 2640
Qy      2641 ATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATG 2700
Db      2641 ATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATG 2700
Qy      2701 GAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGT 2760
Db      2701 GAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGT 2760
Qy      2761 CATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGG 2820
Db      2761 CATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGG 2820
Qy      2821 AAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGA 2880
Db      2821 AAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGA 2880
Qy      2881 AATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAAC 2940
Db      2881 AATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAAC 2940
Qy      2941 TGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAG 3000
Db      2941 TGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAG 3000
Qy      3001 AGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCT 3060
Db      3001 AGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCT 3060
Qy      3061 GGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCAC 3120
Db      3061 GGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCAC 3120
Qy      3121 GCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGG 3180
Db      3121 GCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGG 3180
Qy      3181 GACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACA 3240
Db      3181 GACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACA 3240
Qy      3241 TGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGAT 3300
Db      3241 TGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGAT 3300
Qy      3301 CGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGAC 3360
Db      3301 CGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGAC 3360
Qy      3361 ATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCA 3420
Db      3361 ATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCA 3420
Qy      3421 GCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTAC 3480
Db      3421 GCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTAC 3480
Qy      3481 AAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTT 3540
Db      3481 AAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTT 3540
```

```
Qy       3541 GTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGGTGCAAACTGAACAATGA 3597
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3541 GTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGGTGCAAACTGAACAATGA 3597


RESULT 2
AR164817
LOCUS       AR164817               4072 bp    DNA     linear   PAT 17-OCT-2001
DEFINITION  Sequence 16 from patent US 6274339.
ACCESSION   AR164817
VERSION     AR164817.1  GI:16238027
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 4072)
  AUTHORS   Moore,K. and Nagle,D.Lynn.
  TITLE     Methods and compositions for the diagnosis and treatment of body
            weight disorders, including obesity
  JOURNAL   Patent: US 6274339-A 16 14-AUG-2001;
FEATURES             Location/Qualifiers
     source          1. .4072
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             97.1%;  Score 3494.4;  DB 6;  Length 4072;
  Best Local Similarity   98.8%;  Pred. No. 0;
  Matches 3541;  Conservative    0;  Mismatches   41;  Indels    2;  Gaps    2;

Qy          15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74
               || |||    |  ||   | ||  |  |  |  | | |  |||| ||  || || ||
Db         234 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAA 293

Qy          75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134
               |   |  |  ||  ||||||||||||||||||||||||||||||||||||||||||||
Db         294 GGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 353

Qy         135 CTGCCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         354 CTGCCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 413

Qy         195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         414 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 473

Qy         255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         474 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 533

Qy         315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         534 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 593

Qy         375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         594 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 653

Qy         435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         654 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 713

Qy         495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         714 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 773

Qy         555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         774 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 833

Qy         615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 674
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         834 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 893

Qy         675 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 734
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         894 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 953

Qy         735 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 794
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         954 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 1013

Qy         795 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 854
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1014 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 1073

Qy         855 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 914
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1074 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 1133

Qy         915 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 974
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1134 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 1193

Qy         975 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1034
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1194 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1253

Qy        1035 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1094
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1254 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1313

Qy        1095 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1154
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1314 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1373

Qy        1155 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1214
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1374 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1433

Qy        1215 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1274
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1434 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1493

Qy        1275 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1334
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1494 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1553

Qy        1335 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1394
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1554 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1613

Qy        1395 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1454
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1614 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1673

Qy        1455 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1514
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1674 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1733

Qy        1515 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1574
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1734 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1793

Qy        1575 CTGGTCAGTGCTTCCCAGACCTGAT-TCCACCATGATGTCAACAGATTTGGCCATTCCAG 1633
               ||||||||||||||||||||||||| ||||||||||||||||||||||||||||| |||
Db        1794 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATT-CAG 1852

Qy        1634 CAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCG 1693
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1853 CAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCG 1912

Qy        1694 ACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAG 1753
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1913 ACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAG 1972
```

```
Qy       1754 CAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGG 1813
Db       1973 CAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGG 2032

Qy       1814 CGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTC 1873
Db       2033 CGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTC 2092

Qy       1874 TTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATG 1933
Db       2093 TTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATG 2152

Qy       1934 ACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGA 1993
Db       2153 ACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGA 2212

Qy       1994 TCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGA 2053
Db       2213 TCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGA 2272

Qy       2054 AGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGG 2113
Db       2273 AGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGG 2332

Qy       2114 AGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCAT 2173
Db       2333 AGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCAT 2392

Qy       2174 GTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACC 2233
Db       2393 GTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACC 2452

Qy       2234 ACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGC 2293
Db       2453 ACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGC 2512

Qy       2294 TGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCC 2353
Db       2513 TGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCC 2572

Qy       2354 TTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTT 2413
Db       2573 TTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTT 2632

Qy       2414 TACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAAC 2473
Db       2633 TACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAAC 2692

Qy       2474 CCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTG 2533
Db       2693 CCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTG 2752

Qy       2534 AAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCAT 2593
Db       2753 AAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCAT 2812

Qy       2594 GTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTG 2653
Db       2813 GTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTG 2872

Qy       2654 TGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGA 2713
Db       2873 TGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGA 2932

Qy       2714 TGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGC 2773
Db       2933 TGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGC 2992

Qy       2774 AACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGG 2833
Db       2993 AACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGG 3052

Qy       2834 GTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCAC 2893
Db       3053 GTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCAC 3112

Qy       2894 AGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTC 2953
Db       3113 AGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTC 3172

Qy       2954 ACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGA 3013
Db       3173 ACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGA 3232

Qy       3014 AGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTG 3073
Db       3233 AGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTG 3292

Qy       3074 ATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCA 3133
Db       3293 ATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCA 3352

Qy       3134 ACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGC 3193
Db       3353 ACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGC 3412

Qy       3194 TATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTC 3253
Db       3413 TATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTC 3472

Qy       3254 TTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAG 3313
Db       3473 TTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAG 3532

Qy       3314 CTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATG 3373
Db       3533 CTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATG 3592

Qy       3374 CCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGG 3433
Db       3593 CCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGG 3652

Qy       3434 CTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCT 3493
Db       3653 CTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCT 3712

Qy       3494 CTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTA 3553
Db       3713 CTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTA 3772

Qy       3554 ATTTCACCTGGCCCATCAAAATTCAGGTGCAAACTGAACAATGA 3597
Db       3773 ATTTCACCTGGCCCATCAAAATTCAGGTGCAAACTGAACAATGA 3816
```

RESULT 3
AF106861

```
LOCUS       AF106861               3819 bp    mRNA    linear   PRI 06-JAN-1999
DEFINITION  Homo sapiens attractin-2 (ATRN) mRNA, complete cds.
ACCESSION   AF106861
VERSION     AF106861.1  GI:4093195
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3819)
  AUTHORS   Duke-Cohan,J.S., Gu,J., Freeman,G.J. and Schlossman,S.F.
  TITLE     Cloning of cDNA for attractin-2, identical with that of attractin
            except for a GC-rich 222 bp 5' insertion
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3819)
  AUTHORS   Duke-Cohan,J.S., Gu,J., Freeman,G.J. and Schlossman,S.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-NOV-1998) Division of Tumor Immunology, Dana Farber
            Cancer Institute, 44 Binney Street, Boston, MA 02115, USA
```

```
FEATURES             Location/Qualifiers
     source          1. .3819
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     gene            1. .3819
                     /gene="ATRN"
     CDS             1. .3819
                     /gene="ATRN"
                     /note="similar to Homo sapiens attractin in GenBank
                     Accession Number AF034957"
                     /codon_start=1
                     /product="attractin-2"
                     /protein_id="AAD03057.1"
                     /db_xref="GI:4093196"
                     /translation="MVAAAAATEARLRRRTAATAALAGRSGGPHWDWDVTRAGRPGLG
                     AGLRLPRLLSPPLRPRLLLLLLLLPPPLLLLLLPCEAEAAAAAAAAVSGSAAAEAKECD
                     RPCVNGGRCNPGTGQCVCPAGWVGEQCQHCGGRFRLTGSSGFVTDGPGNYKYKTKCTW
                     LIEGQPNRIMRLRFNHFATECSWDHLYVYDGDSIYAPLVAAFSGLIVPERDGNETVPE
                     VVATSGYALLHFFSDAAYNLTGFNITYSFDMCPNNCSGRGECKISNSSETVECECSEN
                     WKGEACDIPHCTDNCGFPHRGICNSSDVRGCSCFSDWQGPGCSVPVPANQSFWTREEY
                     SNLKLPRASHKAVVNGNIMWVVGGYMFNHSDYNMVLAYDLASREWLPLNRSVNNVVVR
                     YGHSLALYKDKIYMYGGKIDPTGNVTNELRVFHIHNESWVLLTPKAKEQYAVVGHSAH
                     IVTLKNGRVVMLVIFGHCPLYGYISNVQEYDLDKNTWSILHTQGALVQGGYGHSSVYD
                     HRTRALYVHGGYKAFSANKYRLADDLYRYDVDTQMWTILKDSRFFRYLHTAVIVSGTM
                     LVFGGNTHNDTSMSHGAKCFSSDFMAYDIACDRWSVLPRPDLHHDVNRFGHSAVLHNS
                     TMYVFGGFNSLLLSDILVFTSEQCDAHRSEAACLAAGPGIRCVWNTGSSQCISWALAT
                     DEQEEKLKSECFSKRTLDHDRCDQHTDCYSCTANTNDCHWCNDHCVPRNHSCSEGQIS
                     IFRYENCPKDNPMYYCNKKTSCRSCALDQNCQWEPRNQECIALPENICGIGWHLVGNS
                     CLKITTAKENYDNAKLFCRNHNALLASLTTQKKVEFVLKQLRIMQSSQSMSKLTLTPW
                     VGLRKINVSYWCWEDMSPPTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLN
                     GSVCERPANHSAKQCRTPCALRTACGDCTSGSSECMWCSNMKQCVDSNAYVASFPFGQ
                     CMEWYTMSTCPPENCSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKMPSQ
                     APTGNFYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCENLTTGKH
                     CETCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGDECQLCEVENRY
                     QGNPLRGTCYYTLLIDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASKNFN
                     LNITWAASFSAGTQAGEEMPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTW
                     PIKIQVQTEQ"
ORIGIN
```

Query Match 96.9%; Score 3487; DB 9; Length 3819;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 3524; Conservative 0; Mismatches 25; Indels 2; Gaps 2;

```
Qy        48 GGCGGCGACGGCAGCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGTGTCAACGGCGGTCG 107
Db       270 GTCGGGCTCAGCCGCAGCCGAGGCCAAGGAATGTGACCGGCCCTGTGTCAACGGCGGTCG 329

Qy       108 CTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCA 167
Db       330 CTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCA 389

Qy       168 GCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGG 227
Db       390 GCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGG 449

Qy       228 AAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAAT 287
Db       450 AAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAAT 509

Qy       288 GAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGA 347
Db       510 GAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGA 569

Qy       348 TGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAG 407
Db       570 TGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAG 629

Qy       408 AGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTT 467
Db       630 AGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTT 689

Qy       468 TTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTG 527
Db       690 TTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTG 749

Qy       528 TCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGA 587
Db       750 TCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGAAACTGTTGA 809

Qy       588 ATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA 647
Db       810 ATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA 869

Qy       648 CTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTT 707
Db       870 CTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTT 929

Qy       708 CTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGAC 767
Db       930 CTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGAC 989

Qy       768 TCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGG 827
Db       990 TCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGG 1049

Qy       828 AAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCT 887
Db      1050 AAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCT 1109

Qy       888 AGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGT 947
Db      1110 AGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGT 1169

Qy       948 TGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAA 1007
Db      1170 TGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAA 1229

Qy      1008 AATTGATTCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTC 1067
Db      1230 AATTGATCCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTC 1289

Qy      1068 ATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACA 1127
Db      1290 ATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACA 1349

Qy      1128 CATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCT 1187
Db      1350 CATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCT 1409

Qy      1188 CTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATT 1247
Db      1410 CTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATT 1469

Qy      1248 ACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAG 1307
Db      1470 ACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAG 1529

Qy      1308 GACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCT 1367
Db      1530 GACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCT 1589

Qy      1368 TGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAG 1427
Db      1590 TGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAG 1649

Qy      1428 CCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGG 1487
Db      1650 CCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGG 1709

Qy      1488 GGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTT 1547
Db      1710 GGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTT 1769

Qy      1548 CATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGAT-TCCACCA 1606
```

```
Db       1770 CATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCA 1829
Qy       1607 TGATGTCAACAGATTTGGCCATTCCAGCAGTCTTACACAACAGCACCATGTATGTGTTCG 1666
Db       1830 TGATGTCAACAGATTTGGCCATT-CAGCAGTCTTACACAACAGCACCATGTATGTGTTCG 1888
Qy       1667 GTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATG 1726
Db       1889 GTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATG 1948
Qy       1727 CGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACA 1786
Db       1949 CGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACA 2008
Qy       1787 CAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAA 1846
Db       2009 CAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAA 2068
Qy       1847 AATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATT 1906
Db       2069 AATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATT 2128
Qy       1907 GTTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCA 1966
Db       2129 GTTACAGCTGTACAGCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCA 2188
Qy       1967 GGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGG 2026
Db       2189 GGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGG 2248
Qy       2027 ATAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGA 2086
Db       2249 ATAACCCCATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGA 2308
Qy       2087 ACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCA 2146
Db       2309 ACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCA 2368
Qy       2147 TTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATG 2206
Db       2369 TTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATG 2428
Qy       2207 ACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGA 2266
Db       2429 ACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGA 2488
Qy       2267 AGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCA 2326
Db       2489 AGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCA 2548
Qy       2327 AGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGG 2386
Db       2549 AGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGG 2608
Qy       2387 AAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATG 2446
Db       2609 AAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATG 2668
Qy       2447 CTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCA 2506
Db       2669 CTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCA 2728
Qy       2507 TCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCC 2566
Db       2729 TCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCC 2788
Qy       2567 GGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCA 2626
Db       2789 GGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCA 2848
Qy       2627 TGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTT 2686
Db       2849 TGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTT 2908
```

```
Qy       2687 TTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCCTGAAAATTGTTCAGGCT 2746
Db       2909 TTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCCTGAAAATTGTTCAGGCT 2968
Qy       2747 ACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCA 2806
Db       2969 ACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCA 3028
Qy       2807 ATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGC 2866
Db       3029 ATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGC 3088
Qy       2867 AAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGG 2926
Db       3089 AAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGG 3148
Qy       2927 ACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTA 2986
Db       3149 ACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTA 3208
Qy       2987 AATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCG 3046
Db       3209 AATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCG 3268
Qy       3047 AGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCA 3106
Db       3269 AGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCA 3328
Qy       3107 AGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCA 3166
Db       3329 AGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCA 3388
Qy       3167 AGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACC 3226
Db       3389 AGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACC 3448
Qy       3227 CTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTAT 3286
Db       3449 CTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTAT 3508
Qy       3287 CCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAA 3346
Db       3509 CCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAA 3568
Qy       3347 ACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGG 3406
Db       3569 ACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGG 3628
Qy       3407 CTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCA 3466
Db       3629 CTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCA 3688
Qy       3467 ACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAA 3526
Db       3689 ACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAA 3748
Qy       3527 ATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGGTGCAAA 3586
Db       3749 ATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGGTGCAAA 3808
Qy       3587 CTGAACAATGA 3597
Db       3809 CTGAACAATGA 3819
```

```
RESULT 4
AR164816
LOCUS       AR164816               8589 bp    DNA     linear   PAT 17-OCT-2001
DEFINITION  Sequence 14 from patent US 6274339.
ACCESSION   AR164816
VERSION     AR164816.1  GI:16238025
KEYWORDS    .
```

```
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 8589)
  AUTHORS   Moore,K. and Nagle,D.Lynn.
  TITLE     Methods and compositions for the diagnosis and treatment of body
            weight disorders, including obesity
  JOURNAL   Patent: US 6274339-A 14 14-AUG-2001;
FEATURES             Location/Qualifiers
     source          1. .8589
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             96.7%;  Score 3476.8;  DB 6;  Length 8589;
  Best Local Similarity   98.8%;  Pred. No. 0;
  Matches 3524;  Conservative    0;  Mismatches   42;  Indels    2;  Gaps    2;

Qy         15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74
              || |||     |  ||  | ||  |  |  | | |    |||| ||    || || ||
Db        234 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAA 293

Qy         75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134
              |    | |  | ||  ||||||||||||||||||||||||||||||||||||||||||
Db        294 GGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 353

Qy        135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        354 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 413

Qy        195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 473

Qy        255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 533

Qy        315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 593

Qy        375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 653

Qy        435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        654 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 713

Qy        495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        714 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 773

Qy        555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        774 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 833

Qy        615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        834 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 893

Qy        675 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        894 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 953

Qy        735 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        954 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 1013

Qy        795 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1014 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 1073

Qy        855 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1074 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 1133

Qy        915 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1134 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 1193

Qy        975 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1194 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1253

Qy       1035 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1254 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1313

Qy       1095 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1314 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1373

Qy       1155 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1374 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1433

Qy       1215 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1434 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1493

Qy       1275 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1494 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1553

Qy       1335 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1554 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1613

Qy       1395 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1454
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1614 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1673

Qy       1455 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1674 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1733

Qy       1515 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1734 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1793

Qy       1575 CTGGTCAGTGCTTCCCAGACCTGAT-TCCACCATGATGTCAACAGATTTGGCCATTCCAG 1633
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||| |||
Db       1794 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATT-CAG 1852

Qy       1634 CAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCG 1693
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1853 CAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCG 1912

Qy       1694 ACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAG 1753
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1913 ACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAG 1972

Qy       1754 CAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGG 1813
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1973 CAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGG 2032

Qy       1814 CGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTC 1873
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2033 CGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTC 2092

Qy       1874 TTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATG 1933
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2093 TTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATG 2152
```

```
Qy       1934 ACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGA 1993
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2153 ACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGA 2212

Qy       1994 TCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGA 2053
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2213 TCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGA 2272

Qy       2054 AGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGG 2113
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2273 AGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGG 2332

Qy       2114 AGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCAT 2173
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2333 AGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCAT 2392

Qy       2174 GTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACC 2233
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2393 GTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACC 2452

Qy       2234 ACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGC 2293
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2453 ACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGC 2512

Qy       2294 TGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCC 2353
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2513 TGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCC 2572

Qy       2354 TTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTT 2413
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2573 TTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTT 2632

Qy       2414 TACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAAC 2473
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2633 TACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAAC 2692

Qy       2474 CCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTG 2533
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2693 CCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTG 2752

Qy       2534 AAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCAT 2593
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2753 AAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCAT 2812

Qy       2594 GTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTG 2653
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2813 GTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTG 2872

Qy       2654 TGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGA 2713
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2873 TGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGA 2932

Qy       2714 TGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGC 2773
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2933 TGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGC 2992

Qy       2774 AACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGG 2833
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2993 AACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGG 3052

Qy       2834 GTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCAC 2893
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3053 GTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCAC 3112

Qy       2894 AGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTC 2953
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3113 AGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTC 3172

Qy       2954 ACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGA 3013
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3173 ACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGA 3232

Qy       3014 AGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTG 3073
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3233 AGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTG 3292

Qy       3074 ATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCA 3133
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3293 ATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCA 3352

Qy       3134 ACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGC 3193
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3353 ACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGC 3412

Qy       3194 TATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTC 3253
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3413 TATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTC 3472

Qy       3254 TTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAG 3313
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3473 TTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAG 3532

Qy       3314 CTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATG 3373
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3533 CTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATG 3592

Qy       3374 CCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGG 3433
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3593 CCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGG 3652

Qy       3434 CTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCT 3493
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3653 CTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCT 3712

Qy       3494 CTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTA 3553
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3713 CTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTA 3772

Qy       3554 ATTTCACCTGGCCCATCAAAATTCAGGT 3581
             |||||||||||||||||||||||||||||
Db       3773 ATTTCACCTGGCCCATCAAAATTCAGAT 3800
```

RESULT 5
AF531101

```
LOCUS       AF531101                4493 bp    mRNA    linear   MAM 19-AUG-2002
DEFINITION  Bos taurus attractin mRNA, complete cds.
ACCESSION   AF531101
VERSION     AF531101.1  GI:22297307
KEYWORDS    .
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 4493)
  AUTHORS   Graphodatskaya,D., Joerg,H. and Stranzinger,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2002) Animal Sciences, ETH, Tannenstrasse 1,
            Zurich, CH 8092, Switzerland
FEATURES             Location/Qualifiers
     source          1. .4493
                     /organism="Bos taurus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9913"
                     /chromosome="13"
     CDS             52. .4299
                     /note="transmembrane form; mahogany protein"
                     /codon_start=1
                     /product="attractin"
                     /protein_id="AAM95445.1"
                     /db_xref="GI:22297308"
                     /translation="MVAAATEARLRRTVVTPAPTSRNGRRCRDWDATRAGRPGLRTGL
                     RLPRRWRRLLLLLLPPLLLLPWAAEAAAAAAASVSGSAAAEAKECDRPCVNGGRCNPGT
                     GQCVCPAGWVGEQCQHCGGRFRLTGSSGFITDGPGNYKYKTKCTWLIEGQPNRIMRLR
                     FNHFATECSWDHLYVYDGDSIYAPLVAAFSGLIVPERDGNETVPEVVATSGYALLHFF
```

SDAAYNLTGFNITYNFDMCPNNCSGRGECKISNSSNTVQCECSENWKGEACDIPHCVN
NCGFPHRGICNSSDVRGCSCFSEWQGPGCSVPVPANQSFWTREEYSDLKLPRASHKAV
VNGNIMWVVGGYMFNHSDYNMVLAYDLASREWLALNRSVNSVVVRYGHSLALYQDKIY
MYGGKIDSTGNVTNELRVFHIHNESWVLLSPKAKEQYAVVGHSAHIVTLKNGRVVMLV
IFGHCPLYGYISSVQEYDLDKNTWSILHTQGALVQGGYGHSSVYDDRTKALYIHGGYK
AFSANKYRLADDLYRYDVDTQMWTILKDSRFFRYLHTAVIVSGTMLVFGGNTHNDTSM
SHGAKCFSSDFMAYDIDCDRWSVLPRPDLHHDVNRFGHSAVLYNRTMYVFGGFNSLLL
SDILVFTSEQCEAHQSEAACLAAGPGVRCMWDTGSSQCVSWELAPEAQEKIKSECFSK
RIFDHDRCDQLTDCYSCTANTNGCQWCNDHCVPTNHSCTEGQISIFKYDHCPKDNPMY
YCNKKTSCRSCALDQNCQWEPRNQECIALPENICGIGWHLVGNSCLKITTARETYDNA
KLSCRNHNAFLASLTTQKKVEFVLKQLRIMQSSQSMSKFTLTPWVGLRKINVSYWCWE
DMSPFTNSSLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLNGSVCERPANHSAKQ
CRTPCALRTACGECTSGSSECMWCSNMKQCVDSNAYVASFPFGQCMEWYTMSSCPPEN
CSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKMPSQGPTGNSYPQPLLNS
SMCLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCENLTTGKHCETCISGFYGDPTN
GGKCQPCRCNGHASLCNTNTGKCFCTTKGVKGDECQLCEVENRYQGNPLKGTCYYTLL
IDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASKNFNLNITWAASFSAGTQ
AGEEMPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTWPIKIQIAFSQHSNF
MDLVQFFVTFFSCFLSLLLVAAVVWKIKQSCWASRRREQLLREMQQMASRPFASVNVA
LETDEEPPDLIGGSIKTVPKPIALEPCFGNKAAVLSVFVRLPRGLGGIPPPGQSGLAV
ASALVDISQQMPVVYKEKSGAVRNRKQQPPAQPGTCI"

ORIGIN

Query Match 86.2%; Score 3100.8; DB 4; Length 4493;
Best Local Similarity 92.4%; Pred. No. 0;
Matches 3296; Conservative 0; Mismatches 267; Indels 5; Gaps 3;

```
Qy        15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG  74
             ||||||    |  ||   |     |  |  | |||   | ||  ||  || || ||
Db       249 GGCGGCTGAGGCCGCGGCGGCAGCGGCGTCGGTGTCGGGCTCAGCCGCAGCCGAGGCCAA 308

Qy        75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134
             |    | |  | || |||||||||||||||| |||||||||||||||||||||||||| ||
Db       309 GGAATGTGACCGGCCCTGTGTCAACGGCGGCCGCTGCAACCCTGGCACCGGCCAGTGTGT 368

Qy       135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       369 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 428

Qy       195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254
             |||||||||||| ||| | ||||||||||||||||||||||||| ||||| || |||||
Db       429 TGGATCTTCTGGATTTATAACAGATGGACCTGGAAATTATAAATATAAAACAAAATGCAC 488

Qy       255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       489 ATGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 548

Qy       315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       549 AGAGTGTAGTTGGGACCATTTATATGTTTATGACGGGGACTCAATTTATGCACCCCTAGT 608

Qy       375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434
             |||||| ||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db       609 TGCTGCCTTTAGCGGCCTCATTGTTCCTGAGAGAGACGGCAATGAGACTGTCCCAGAGGT 668

Qy       435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494
             ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       669 TGTCGCCACATCAGGTTATGCCCTGCTGCATTTTTTTAGTGATGCTGCTTATAACTTGAC 728

Qy       495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554
             |||||||||||| ||||||| |||||| |||||||||||||| ||||||||||||||||||
Db       729 TGGATTTAATATCACTTACAATTTTGACATGTGTCCAAATAATTGCTCAGGCCGAGGAGA 788

Qy       555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614
             |||||||||||||||||||| | |||||| ||||||||||||||||||||||||||||||
Db       789 ATGTAAGATCAGTAATAGCAGCAACACTGTTCAATGTGAATGTTCTGAAAACTGGAAAGG 848

Qy       615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 674
             |||||||||||||||||||||| | |||||||||||||||||||||||||||||||||||
Db       849 TGAAGCATGTGACATTCCTCACTGTGTAAATAACTGTGGTTTTCCTCATCGAGGCATCTG 908

Qy       675 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 734
             |||||||||||||||||||||||||||||||||||||||||| ||||| ||||||||||||
```

```
Db       909 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGAGTGGCAAGGTCCTGGATGTTC 968

Qy       735 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 794
             |||||||||||||||||||||||||||||||||||||||| |||||||||||||| || ||||||
Db       969 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACACGAGAGGAATATTCTGACCTAAAGCT 1028

Qy       795 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 854
             ||||||||| |||| |||||||||||||||||||||| || ||||||||||||| |||||
Db      1029 CCCCAGAGCCTCTCACAAAGCTGTGGTCAATGGAAATATAATGTGGGTTGTTGGCGGATA 1088

Qy       855 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 914
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db      1089 TATGTTCAACCACTCAGATTACAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 1148

Qy       915 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 974
             |||| |||| |||||||||||||||| |||||||||||||||||| |||||||| ||||
Db      1149 GCTTGCACTGAACCGTTCTGTGAACAGTGTGGTTGTTAGATATGGCCATTCTTTAGCATT 1208

Qy       975 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1034
             |||| |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db      1209 ATACCAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACAGGGAATGTGACCAA 1268

Qy      1035 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1094
             |||||||||||| |||| |||||||||||| |||||||||||||||| |||| || ||||
Db      1269 TGAGTTGAGAGTGTTTCATATTCATAATGAATCATGGGTGTTGTTGTCCCCCAAAGCAAA 1328

Qy      1095 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1154
             |||||||||||||||||||||||||||||| || |||||||| || |||| ||||||||||
Db      1329 GGAGCAGTATGCAGTGGTTGGGCACTCGGCGCACATTGTCACGTTGAAAAATGGCCGAGT 1388

Qy      1155 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1214
             |||||||||||| ||||||||||||||||||||||||||||||||||||| |||||||||
Db      1389 GGTCATGCTGGTTATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAGTGTGCAGGA 1448

Qy      1215 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1274
             |||||| ||||||||||||||||||||||||||||||||| |||||||||||||| |||||
Db      1449 ATATGACTTGGATAAGAACACATGGAGTATATTACACACTCAGGGTGCCCTTGTCCAAGG 1508

Qy      1275 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1334
             || |||||||||||||||||||||| || |||||||||| || |||| || ||| || ||
Db      1509 CGGCTACGGCCATAGCAGTGTCTATGATGATAGGACCAAGGCCCTGTATATTCACGGAGG 1568

Qy      1335 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1394
             |||||||||||||||||||||||| |||||||||||||||||||||| ||||||||||||
Db      1569 CTACAAGGCTTTCAGTGCCAATAAATACCGGCTTGCAGATGATCTCTACAGATATGATGT 1628

Qy      1395 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1454
             ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1629 GGACACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1688

Qy      1455 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1514
             |||||||||||||||||||||||||||||||||||||||||||||||||||| || ||||
Db      1689 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGAGGAAACACACACAATGACACGTCCAT 1748

Qy      1515 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1574
             |||||||| |||||||||||||||||||||||||||||||| |||||||| ||||||||
Db      1749 GAGCCATGGGGCCAAATGCTTCTCTTCAGATTTCATGGCTTATGACATTGATTGTGACCG 1808

Qy      1575 CTGGTCAGTGCTTCCCAGACCTGAT-TCCACCATGATGTCAACAGATTTGGCCATTCCAG 1633
             |||||||||||||||||||||||| ||||||||||||||||||||||||||| | |||
Db      1809 ATGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCACT-CAG 1867

Qy      1634 CAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCG 1693
             | |||||| ||||||| |||||||||| || |||||||| |||||||||||||||||| |
Db      1868 CGGTCTTATACAACAGAACCATGTATGTATTTGGTGGTTTTAATAGTCTCCTCCTCAGTG 1927

Qy      1694 ACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAG 1753
             ||||||| |||||||| |||||||||||||| || || | |||||||| |||| | |
Db      1928 ACATCCTAGTATTCACATCGGAACAGTGTGAAGCACACCAGAGTGAAGCTGCTTGTCTCG 1987

Qy      1754 CAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGG 1813
             |||||||||||| |||||||| |||| |||||||||| |||||||| |||| ||||
Db      1988 CAGCAGGACCTGGTGTTCGGTGTATGTGGGACACAGGGTCATCTCAGTGTGTCTCCTGGG 2047
```

```
Qy       1814 CGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTC 1873
Db       2048 AGCTGGCACCTGAAGCAC---AAGAAAAGATAAAGTCAGAATGTTTTTCTAAAAGAATTT 2104

Qy       1874 TTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATG 1933
Db       2105 TTGACCATGACAGATGTGACCAGCTCACGGATTGTTACAGCTGCACAGCCAACACCAATG 2164

Qy       1934 ACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGA 1993
Db       2165 GCTGCCAGTGGTGCAATGACCATTGTGTCCCTACGAACCACAGCTGCACAGAAGGCCAGA 2224

Qy       1994 TCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGA 2053
Db       2225 TCTCCATTTTCAAGTATGATCATTGCCCCAAGGATAACCCGATGTATTATTGTAACAAGA 2284

Qy       2054 AGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGG 2113
Db       2285 AGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGG 2344

Qy       2114 AGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCAT 2173
Db       2345 AGTGCATCGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCAT 2404

Qy       2174 GTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACC 2233
Db       2405 GTTTGAAAATTACTACTGCCAGGGAGACTTATGACAATGCAAAATTGTCCTGCAGAAACC 2464

Qy       2234 ACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGC 2293
Db       2465 ACAATGCCTTTTTGGCTTCTCTTACAACTCAGAAGAAGGTAGAATTTGTCCTTAAGCAGC 2524

Qy       2294 TGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCC 2353
Db       2525 TGCGAATCATGCAGTCATCACAGAGCATGTCCAAGTTCACCTTAACCCCATGGGTTGGCC 2584

Qy       2354 TTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTT 2413
Db       2585 TTCGGAAGATCAATGTGTCCTATTGGTGCTGGGAAGATATGTCCCCGTTCACAAATAGTT 2644

Qy       2414 TACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAAC 2473
Db       2645 CACTACAGTGGATGCCATCTGAGCCCAGTGATGCTGGATTCTGTGGGATCTTGTCAGAAC 2704

Qy       2474 CCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTG 2533
Db       2705 CCAGTACTCGGGGCTTAAAGGCTGCAACCTGCATCAATCCACTCAATGGTAGTGTCTGTG 2764

Qy       2534 AAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCAT 2593
Db       2765 AAAGGCCTGCTAACCACAGTGCCAAGCAGTGTCGGACCCCTTGTGCCTTGCGGACAGCCT 2824

Qy       2594 GTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTG 2653
Db       2825 GTGGGGAGTGTACCAGTGGCAGCTCGGAGTGCATGTGGTGCAGCAACATGAAGCAGTGCG 2884

Qy       2654 TGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGA 2713
Db       2885 TGGACTCCAATGCCTACGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAGTGGTATACCA 2944

Qy       2714 TGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGC 2773
Db       2945 TGAGCAGCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCACTGCTTGGAGC 3004

Qy       2774 AACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGG 2833
Db       3005 AACCAGGCTGTGGCTGGTGCACTGATCCTAGCAATACGGGCAAAGGGAAATGCATAGAGG 3064

Qy       2834 GTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCAC 2893
Db       3065 GCTCCTATAAAGGACCAGTGAAGATGCCTTCTCAGGGCCCTACAGGAAATTCCTACCCAC 3124

Qy       2894 AGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTC 2953
Db       3125 AGCCTCTTCTCAATTCCAGCATGTGTCTGGAGGACAGCAGATATAACTGGTCTTTCATTC 3184

Qy       2954 ACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGA 3013
Db       3185 ATTGTCCAGCTTGCCAGTGCAATGGCCACAGCAAATGCATTAATCAGAGTATCTGTGAGA 3244

Qy       3014 AGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTG 3073
Db       3245 AGTGTGAGAACCTGACCACGGGCAAGCACTGTGAGACCTGCATATCTGGCTTCTATGGTG 3304

Qy       3074 ATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCA 3133
Db       3305 ATCCAACCAACGGGGGAAAATGTCAACCATGCAGGTGCAATGGACATGCGTCGCTATGCA 3364

Qy       3134 ACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGC 3193
Db       3365 ATACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGATGAATGTCAGC 3424

Qy       3194 TATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTC 3253
Db       3425 TATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAAAGGAACGTGTTATTATACTC 3484

Qy       3254 TTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAG 3313
Db       3485 TTCTTATTGACTATCAGTTCACCTTTAGCCTATCCCAGGAAGATGACCGCTATTACACAG 3544

Qy       3314 CTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATG 3373
Db       3545 CCATCAATTTTGTGGCGACACCTGATGAACAAAACAGGGATTTGGACATGTTCATCAATG 3604

Qy       3374 CCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGG 3433
Db       3605 CCTCCAAAAACTTCAACCTCAACATCACCTGGGCTGCCAGCTTCTCAGCTGGAACCCAAG 3664

Qy       3434 CTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCT 3493
Db       3665 CTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCT 3724

Qy       3494 CTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTA 3553
Db       3725 CTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTA 3784

Qy       3554 ATTTCACCTGGCCCATCAAAATTCAGGT 3581
Db       3785 ATTTCACCTGGCCCATCAAAATTCAGAT 3812
```

```
RESULT 6
AF116897
LOCUS       AF116897               6083 bp    mRNA    linear   ROD 20-MAR-1999
DEFINITION  Mus musculus mahogany protein mRNA, complete cds.
ACCESSION   AF116897
VERSION     AF116897.1  GI:4454560
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 6083)
  AUTHORS   Nagle,D.L., McGrail,S.H., Vitale,J., Woolf,E.A., Dussault,B.J. Jr.,
            DiRocco,L., Holmgren,L., Montagno,J., Bork,P., Huszar,D.,
            Fairchild-Huntress,V., Ge,P., Keilty,J., Ebeling,C., Baldini,L.,
            Gilchrist,J., Burn,P., Carlson,G.A. and Moore,K.J.
  TITLE     The mahogany protein is a receptor involved in suppression of
            obesity
  JOURNAL   Nature 398 (6723), 148-152 (1999)
  MEDLINE   99184159
   PUBMED   10086355
REFERENCE   2  (bases 1 to 6083)
  AUTHORS   Nagle,D.L., McGrail,S.H., Vitale,J., Woolf,E.A., Dussault,B.J. Jr.,
            DiRocco,L., Holmgren,L., Montagno,J., Bork,P., Huszar,D.,
```

```
            Fairchild-Huntress,V., Ge,P., Keilty,J., Ebeling,C., Baldini,L.,
            Gilchrist,J., Burn,P., Carlson,G.A. and Moore,K.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-DEC-1998) Sequence Analysis, Millennium
            Pharmaceuticals, 640 Memorial Drive, Cambridge, MA 02139, USA
FEATURES             Location/Qualifiers
     source          1. .6083
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
     CDS             70. .4488
                     /note="MG"
                     /codon_start=1
                     /product="mahogany protein"
                     /protein_id="AAD20947.1"
                     /db_xref="GI:4454561"
                     /translation="MVAVAAAAATEARLRGSTTTTAAPAGRKGRQHRPCTATGAWRPG
                     PRARLCLPRVLSRALPPPPLLPLLFSLLLLPLPREAEAAAVAAAVSGSAAAEAKECDR
                     PCVNGGRCNPGTGQCVCPTGWVGEQCQHCGGRFRLTGSSGFVTDGPGNYKYKTKCTWL
                     IEGQPNRIMRLRFNHFATECSWDHLYVYDGDSIYAPLIAAFSGLIVPERDGNETAPEV
                     TVTSGYALLHFFSDAAYNLTGFNITYNFDMCPNNCSGRGECKSSNSSSAVECECSENW
                     KGESCDIPHCTDNCGFPHRGICNASDTRGCSCFPHWQGPGCSIPVPANQSFWTREEYS
                     DLKLPRASHKAVVNGNIMWVVGGYMFNHSDYSMVLAYDLTSREWLPLNHSVNSVVVRY
                     GHSLALHKDKIYMYGGKIDSTGNVTNELRVFHIHNESWVLLTPKAKDQYAVVGHSAHI
                     VTLASGRVVMLVIFGHCPLYGYISVVQEYDLEKNTWSILHTQGALVQGGYGHSSVYDD
                     RTKALYVHGGYKAFSANKYRLADDLYRYDVDTQMWTILKDSRFFRYLHTAVIVSGTML
                     VFGGNTHNDTSMSHGAKCFSSDFMAYDIACDRWSVLPRPELHHDVNRFGHSAVLYNST
                     MYVFGGFNSLLLSDVLVFTSEQCDAHRSEAACVAAGPGIRCLWDTQSSRCTSWELATE
                     EQAEKLKSECFSKRTLDHDRCDQHTDCYSCTANTNDCHWCNDHCVPVNHSCTEGQISI
                     AKYESCPKDNPMYYCNKKTSCRSCALDQNCQWEPRNQECIALPENICGNGWHLVGNSC
                     LKITTAKENYDNAKLSCRNHNAFLASLTSQKKVEFVLKQLRLMQSSQSMSKLTLTPWV
                     GLRKINVSYWCWEDMSPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLNG
                     SVCERPANHSAKQCRTPCALRTACGECTSSSSECMWCSNMKQCVDSNAYVASFPFGQC
                     MEWYTMSSCPPENCSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKMPSQA
                     SAGNVYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCEDLTTGKHC
                     ETCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGDECQLCEVENRYQ
                     GNPLKGTCYYTLLIDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASKNFNL
                     NITWATSFPAGTQTGEEVPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTWP
                     IKIQIAFSQHSNFMDLVQFFVTFFSCFLSLLLVAAVVWKIKQSCWASRRREQLLREMQ
                     QMASRPFASVNVALETDEEPPDLIGGSIKTVPKPIALEPCFGNKAAVLSVFVRLPRGL
                     GGIPPPGQSRSRCGQCPGGHFSADANSVQGEVRSCKKPEAAAACTAWNLHLILGQEFA
                     LTEGVWSTKLSTGEAAGRKQAWGRRLGTLEASNSCA"
ORIGIN

  Query Match          82.0%;  Score 2948.8;  DB 10;  Length 6083;
  Best Local Similarity  89.5%;  Pred. No. 0;
  Matches 3194;  Conservative    0;  Mismatches  372;  Indels    2;  Gaps    2;

Qy        15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74
Db       303 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAA 362

Qy        75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134
Db       363 GGAATGTGACCGGCCGTGTGTCAACGGCGGCCGCTGCAACCCTGGCACCGGCCAGTGCGT 422

Qy       135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194
Db       423 CTGCCCCACGGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 482

Qy       195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254
Db       483 TGGCTCTTCTGGATTTGTAACAGATGGACCTGGGAATTATAAATATAAGACGAAGTGCAC 542

Qy       255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314
Db       543 ATGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGCTTCAACCATTTTGCTAC 602

Qy       315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374
Db       603 AGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATCTACGCACCTCTGAT 662

Qy       375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434

Db       663 TGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCTCCTGAGGT 722

Qy       435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494
Db       723 CACTGTCACTTCAGGTTATGCACTGCTGCATTTTTTCAGTGATGCTGCTTATAATCTGAC 782

Qy       495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554
Db       783 TGGATTTAATATCACTTACAATTTTGACATGTGTCCGAATAATTGCTCAGGCCGAGGAGA 842

Qy       555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614
Db       843 GTGTAAGAGCAGTAACAGCAGCAGCGCTGTTGAGTGTGAATGTTCTGAAAACTGGAAAGG 902

Qy       615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 674
Db       903 GGAGTCGTGTGACATTCCTCACTGTACAGACAACTGTGGCTTTCCTCACCGAGGCATCTG 962

Qy       675 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 734
Db       963 TAATGCAAGCGATACCAGAGGGTGCTCCTGCTTTCCTCACTGGCAGGGTCCTGGATGTTC 1022

Qy       735 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 794
Db      1023 AATTCCTGTGCCAGCTAACCAGTCTTTTTGGACTCGAGAAGAATATTCTGATTTAAAGCT 1082

Qy       795 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 854
Db      1083 TCCCAGAGCCTCTCATAAAGCTGTGGTCAATGGAAATATAATGTGGGTTGTTGGCGGATA 1142

Qy       855 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 914
Db      1143 TATGTTCAACCATTCAGATTACAGCATGGTTTTAGCGTATGACCTGACTTCTAGGGAATG 1202

Qy       915 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 974
Db      1203 GCTTCCACTAAACCATTCTGTGAACAGTGTGGTTGTAAGATATGGTCATTCTTTGGCATT 1262

Qy       975 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1034
Db      1263 ACATAAGGATAAAATCTACATGTATGGAGGAAAAATTGATTCAACAGGGAACGTGACCAA 1322

Qy      1035 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1094
Db      1323 TGAGCTGAGAGTATTTCATATTCATAATGAATCATGGGTATTGTTAACTCCGAAAGCTAA 1382

Qy      1095 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1154
Db      1383 GGATCAGTATGCAGTGGTTGGACACTCAGCACACATTGTTACACTGGCATCTGGCCGTGT 1442

Qy      1155 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1214
Db      1443 GGTCATGTTGGTCATCTTCGGTCATTGCCCACTCTATGGATATATAAGCGTTGTGCAGGA 1502

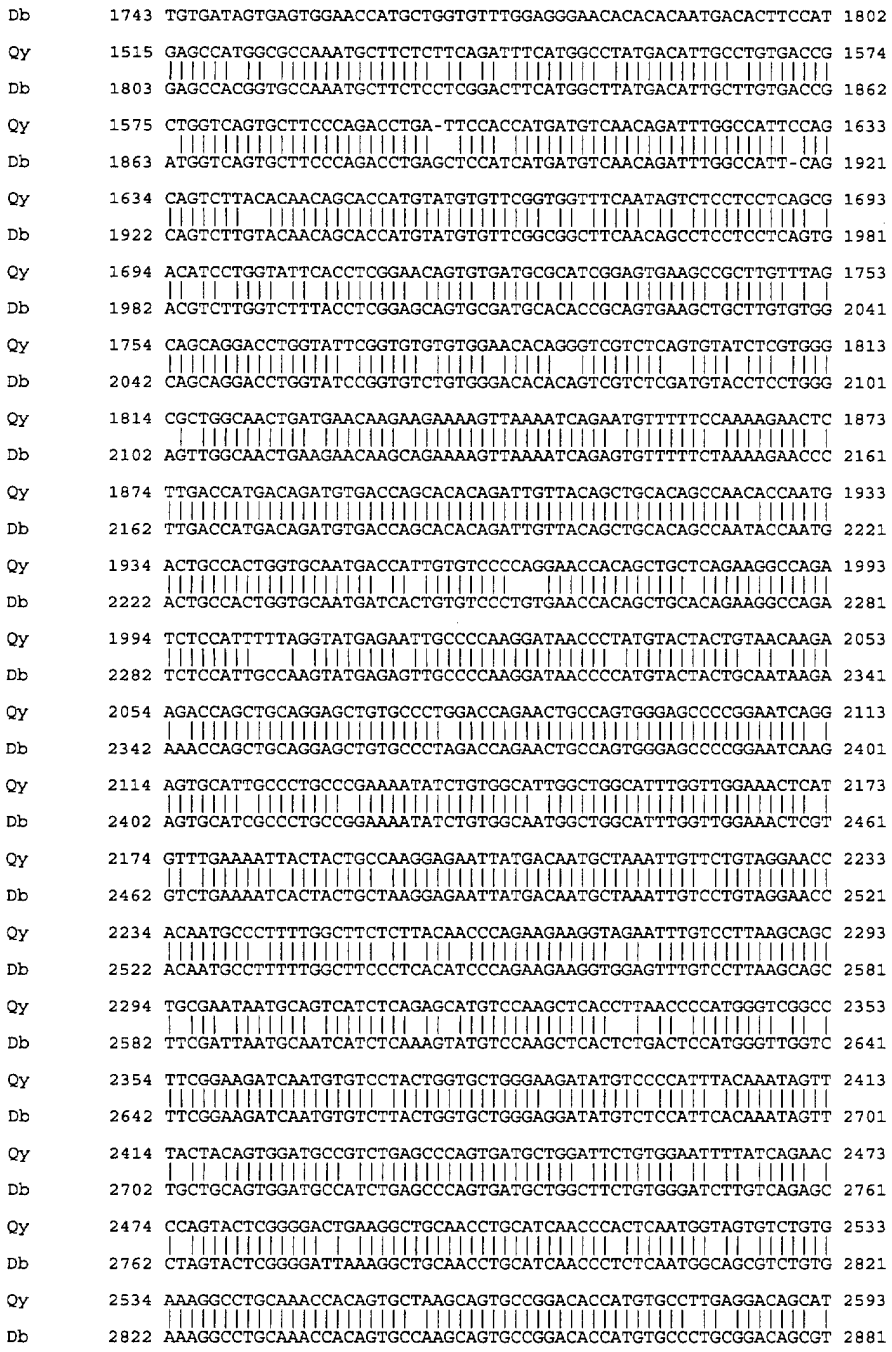
Qy      1215 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1274
Db      1503 ATATGACTTGGAAAAGAACACATGGAGTATATTACATACTCAGGGTGCTCTTGTGCAAGG 1562

Qy      1275 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1334
Db      1563 GGGTTATGGCCACAGTAGTGTTTATGATGACAGGACCAAGGCTCTGTACGTTCATGGTGG 1622

Qy      1335 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1394
Db      1623 CTACAAGGCTTTCAGCGCCAACAAATACCGGCTTGCAGATGACCTCTACAGATACGATGT 1682

Qy      1395 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1454
Db      1683 GGATACTCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCATACAGC 1742

Qy      1455 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1514
```



RESULT 7
AB062913
LOCUS AB062913 4284 bp mRNA linear ROD 16-JAN-2002

```
DEFINITION  Mesocricetus auratus Atrn mRNA for attractin, complete cds.
ACCESSION   AB062913
VERSION     AB062913.1  GI:16930100
KEYWORDS    .
SOURCE      Mesocricetus auratus (golden hamster)
  ORGANISM  Mesocricetus auratus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
            Mesocricetus.
REFERENCE   1
  AUTHORS   Kuramoto,T., Nomoto,T., Fujiwara,A., Mizutani,M., Sugimura,T. and
            Ushijima,T.
  TITLE     Insertional mutation of the Attractin gene in the black tremor
            hamster
  JOURNAL   Mamm. Genome 13 (1), 36-40 (2002)
  MEDLINE   21635551
   PUBMED   11773967
REFERENCE   2  (bases 1 to 4284)
  AUTHORS   Kuramoto,T. and Ushijima,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-JUN-2001) Takashi Kuramoto, National Cancer Center
            Research Institute, Carcinogenesis Division; Tsukiji 5-1-1,
            Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramot@gan2.ncc.go.jp,
            URL:http://www.ncc.go.jp/research/rat-genome/, Tel:81-3-3542-2511,
            Fax:81-3-5565-1753)
FEATURES             Location/Qualifiers
     source          1. .4284
                     /organism="Mesocricetus auratus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10036"
     gene            1. .4284
                     /gene="Atrn"
     CDS             1. .4284
                     /gene="Atrn"
                     /codon_start=1
                     /product="attractin"
                     /protein_id="BAB72012.1"
                     /db_xref="GI:16930101"
                     /translation="MVAAAAATEALLRGSTAATAAPAGRKVEQYRLCAATGAWRPGPR
                     AGLCLPWVVSRALPPPLLLLLLLFSLLLLPLPREAEAAAVSAAVSGSAAAEAKECDRP
                     CVNGGRCNPGTGQCVCPAGWVGEQCQHCGGRFRLTGSSGFVTDGPGNYKYKTKCTWLI
                     EGQPNRIMRLRFNHFATECSWDHLYVYDGDSIYAPLIAAFSGLIVPERDGNETAPEVT
                     VTSGYALLHFFSDAAYNLTGFNITYNFDMCPNNCSGRGECKSSNSSNTVECECSENWK
                     GESCDIPHCTDNCGFPHRGICNASDVRGCSCSPHWQGPGCSIPVPANQSFWIREEYSD
                     FKLPRASHKAVVNGNIMWVVGGYMFNHSDYSMVLAYDLASREWLPLNHSVNNVVVRYG
                     HSLALHKDKIYMYGGKIDSTGNVTNELRVFHIHNESWVLLTPKAKEQYAVVGHSAHIV
                     TLATGRVVMLVIFGHCPLYGYISIVQEYDLEKNTWSILQTQGALVQGGYGHSSVYDHR
                     TKALYVHGGYKAFSANKYRLADDLYRYDVDTQMWTILKDSRFFRYLHTAVIVSGTMLV
                     FGGNTHNDTSMSHGAKCFSSDFMAYDIACDRWSVLPRPDLHHDVNRFGHSAVLHNSTM
                     YVFGGFNSLLLSDVLVFTSEQCDAHHSEAACVAAGPGIRCLWDTVSSRCTSWELATEE
                     QAEKLKSECFSKRTLDHDRCDQHTDCYSCTANTNDCHWCNDHCVPVNHSCTEGQISIS
                     KYDNCPKDNPMYYCNKKTSCRSCALDQNCQWEPRNQECIALPENICGIGWHLVGNSCL
                     KITTAKENYDNAKLSCRNHNAFLASLTSQKKVEFVLKHLRLMQSSQSMSKLTLTPWVG
                     LRKINVSYWCWEDMSPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLNGS
                     VCERPANHSAKQCRTPCALRTACSECTSSSSECMWCSNMKQCVDSNAYVASFPFGQCM
                     EWYTMSSCPPENCSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPIRMPSQAS
                     TGNVYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCENLTTGKHCE
                     TCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGDECQLCEVENRYQG
                     NPLKGTCYYTLLIDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASKNFNLN
                     ITWATSFSAGTQAGEEMPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTWPI
                     KIQIAFSQHSNFMDLVQFFVTFFSCFLSLLLVAAVVWKIKQSCWASRRREQLLREMQQ
                     MASRPFASVNVALETDEEPPDLIGGSIKTVPKPIALEPCFGNKAAVLSVFVRLPRGLG
                     GIPPPGQSGLAVASALVDISQQMPIVYKEKSGAVRNRKQQPPAQPGTCI"
ORIGIN

  Query Match             82.0%;  Score 2947.8;  DB 10;  Length 4284;
  Best Local Similarity   89.5%;  Pred. No. 0;
  Matches 3193;  Conservative    0;  Mismatches  372;  Indels    2;  Gaps    2;

Qy         16 GCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAGG 75
Db        232 GAGGCCGAGGCCGCGGCGGTGTCGGCGGCGGTGTCAGGCTCGGCCGCAGCCGAGGCCAAG 291

Qy         76 AGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTC 135
Db        292 GAATGTGACCGGCCGTGTGTCAACGGCGGCCGCTGCAACCCTGGCACTGGCCAGTGCGTC 351

Qy        136 TGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACT 195
Db        352 TGCCCCGCGGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACT 411

Qy        196 GGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCACG 255
Db        412 GGATCTTCTGGATTCGTAACAGATGGACCTGGGAATTATAAGTATAAAACGAAGTGCACA 471

Qy        256 TGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACA 315
Db        472 TGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACA 531

Qy        316 GAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTT 375
Db        532 GAATGTAGCTGGGACCATTTATATGTTTATGACGGGGACTCAATTTATGCACCTCTGATT 591

Qy        376 GCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTT 435
Db        592 GCTGCCTTTAGTGGTCTCATTGTTCCTGAGAGAGATGGCAATGAGACAGCCCCAGAGGTC 651

Qy        436 GTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACT 495
Db        652 ACTGTCACTTCAGGTTATGCACTGCTGCATTTTTTCAGTGATGCTGCTTACAATCTGACT 711

Qy        496 GGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAG 555
Db        712 GGATTTAATATCACTTACAATTTTGACATGTGTCCGAATAATTGCTCAGGCCGAGGAGAA 771

Qy        556 TGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGT 615
Db        772 TGTAAGAGCAGTAACAGCAGCAATACTGTTGAGTGTGAATGTTCTGAAAACTGGAAAGGG 831

Qy        616 GAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGC 675
Db        832 GAATCGTGTGACATTCCCCATTGTACAGACAACTGTGGGTTTCCTCACCGAGGCATCTGT 891

Qy        676 AATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCA 735
Db        892 AATGCAAGTGACGTGAGAGGGTGTTCCTGCTCCCCGCACTGGCAGGGTCCTGGATGTTCA 951

Qy        736 GTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTC 795
Db        952 ATTCCTGTGCCAGCTAACCAGTCTTTCTGGATTCGAGAAGAATATTCCGATTTCAAGCTC 1011

Qy        796 CCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATAT 855
Db       1012 CCCAGAGCCTCTCATAAAGCTGTGGTCAATGGGAATATAATGTGGGTTGTTGGTGGATAT 1071

Qy        856 ATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGG 915
Db       1072 ATGTTCAACCACTCAGATTACAGCATGGTTTTAGCGTATGACCTGGCTTCTAGGGAATGG 1131

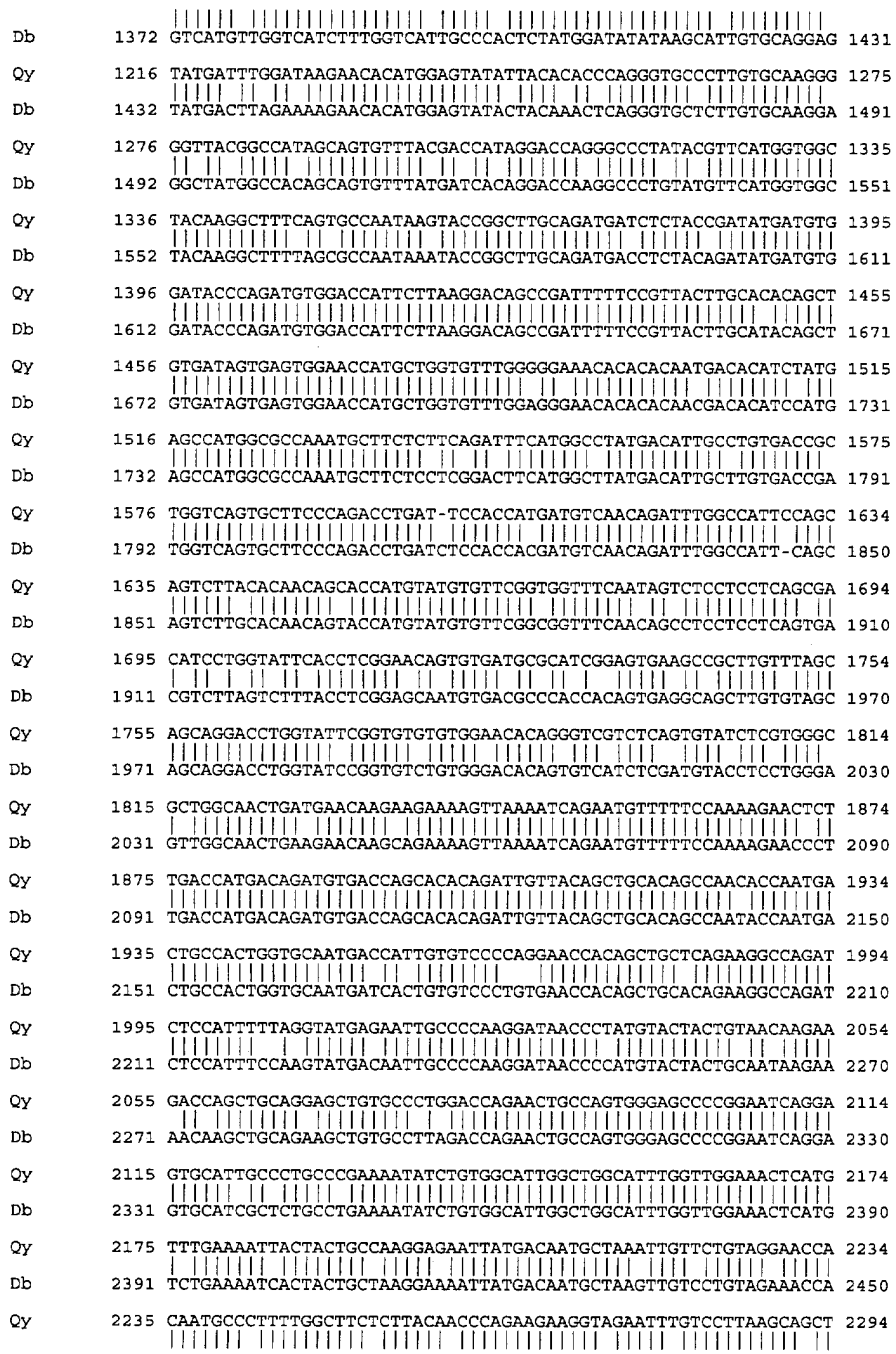
Qy        916 CTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTA 975
Db       1132 CTTCCGCTAAACCATTCTGTGAACAATGTGGTTGTAAGATATGGTCATTCTTTGGCATTG 1191

Qy        976 TACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAAT 1035
Db       1192 CATAAGGATAAAATCTACATGTATGGAGGAAAAATTGATTCAACAGGGAATGTGACCAAT 1251

Qy       1036 GAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAG 1095
Db       1252 GAGCTGAGAGTATTTCATATTCATAATGAATCATGGGTATTGTTAACACCAAAAGCAAAG 1311

Qy       1096 GAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTG 1155
Db       1312 GAGCAGTATGCAGTAGTTGGACACTCAGCACATATTGTTACACTGGCAACCGGCCGTGTG 1371

Qy       1156 GTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAA 1215
```


Db 1372 GTCATGTTGGTCATCTTTGGTCATTGCCCACTCTATGGATATATAAGCATTGTGCAGGAG 1431
Qy 1216 TATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGG 1275
Db 1432 TATGACTTAGAAAAGAACACATGGAGTATACTACAAACTCAGGGTGCTCTTGTGCAAGGA 1491
Qy 1276 GGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGC 1335
Db 1492 GGCTATGGCCACAGCAGTGTTTATGATCACAGGACCAAGGCCCTGTATGTTCATGGTGGC 1551
Qy 1336 TACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTG 1395
Db 1552 TACAAGGCTTTTAGCGCCAATAAATACCGGCTTGCAGATGACCTCTACAGATATGATGTG 1611
Qy 1396 GATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCT 1455
Db 1612 GATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCATACAGCT 1671
Qy 1456 GTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATG 1515
Db 1672 GTGATAGTGAGTGGAACCATGCTGGTGTTTGGAGGGAACACACACAACGACACATCCATG 1731
Qy 1516 AGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGC 1575
Db 1732 AGCCATGGCGCCAAATGCTTCTCCTCGGACTTCATGGCTTATGACATTGCTTGTGACCGA 1791
Qy 1576 TGGTCAGTGCTTCCCAGACCTGAT-TCCACCATGATGTCAACAGATTTGGCCATTCCAGC 1634
Db 1792 TGGTCAGTGCTTCCCAGACCTGATCTCCACCACGATGTCAACAGATTTGGCCATT-CAGC 1850
Qy 1635 AGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGA 1694
Db 1851 AGTCTTGCACAACAGTACCATGTATGTGTTCGGCGGTTTCAACAGCCTCCTCCTCAGTGA 1910
Qy 1695 CATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGC 1754
Db 1911 CGTCTTAGTCTTTACCTCGGAGCAATGTGACGCCCACCACAGTGAGGCAGCTTGTGTAGC 1970
Qy 1755 AGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGC 1814
Db 1971 AGCAGGACCTGGTATCCGGTGTCTGTGGGACACAGTGTCATCTCGATGTACCTCCTGGGA 2030
Qy 1815 GCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCT 1874
Db 2031 GTTGGCAACTGAAGAACAAGCAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACCCT 2090
Qy 1875 TGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGA 1934
Db 2091 TGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAATACCAATGA 2150
Qy 1935 CTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGAT 1994
Db 2151 CTGCCACTGGTGCAATGATCACTGTGTCCCTGTGAACCACAGCTGCACAGAAGGCCAGAT 2210
Qy 1995 CTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAA 2054
Db 2211 CTCCATTTCCAAGTATGACAATTGCCCCAAGGATAACCCCATGTACTACTGCAATAAGAA 2270
Qy 2055 GACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGA 2114
Db 2271 AACAAGCTGCAGAAGCTGTGCCTTAGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGA 2330
Qy 2115 GTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATG 2174
Db 2331 GTGCATCGCTCTGCCTGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATG 2390
Qy 2175 TTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCA 2234
Db 2391 TCTGAAAATCACTACTGCTAAGGAAAATTATGACAATGCTAAGTTGTCCTGTAGAAACCA 2450
Qy 2235 CAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCT 2294
Db 2451 CAATGCCTTTTTGGCTTCCCTTACATCCCAGAAGAAGGTGGAATTCGTCCTTAAGCATCT 2510
Qy 2295 GCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCT 2354
Db 2511 TCGATTAATGCAGTCATCTCAGAGCATGTCCAAGCTCACTCTGACTCCATGGGTTGGTCT 2570
Qy 2355 TCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTT 2414
Db 2571 TCGGAAGATCAATGTGTCTTATTGGTGCTGGGAGGATATGTCTCCATTTACAAATAGTTT 2630
Qy 2415 ACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACC 2474
Db 2631 GCTGCAGTGGATGCCATCCGAGCCCAGTGATGCTGGTTTCTGTGGGATCTTGTCAGAGCC 2690
Qy 2475 CAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGA 2534
Db 2691 TAGTACTCGGGGATTGAAGGCTGCAACCTGCATCAACCCACTCAATGGCAGTGTTTGTGA 2750
Qy 2535 AAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATG 2594
Db 2751 AAGGCCTGCAAACCACAGTGCCAAGCAGTGCCGGACACCGTGTGCCCTGCGGACAGCGTG 2810
Qy 2595 TGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGT 2654
Db 2811 TAGTGAGTGCACTAGCAGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGT 2870
Qy 2655 GGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGAT 2714
Db 2871 AGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAGTGGTATACAAT 2930
Qy 2715 GAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCA 2774
Db 2931 GAGCAGCTGCCCACCTGAAAATTGCTCTGGCTACTGTACCTGCAGCCATTGCTTGGAGCA 2990
Qy 2775 ACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGG 2834
Db 2991 GCCAGGCTGTGGCTGGTGTACTGATCCTAGCAATACTGGGAAAGGGAAATGTATCGAGGG 3050
Qy 2835 TTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACA 2894
Db 3051 CAGCTATAAAGGACCCATAAGGATGCCTTCACAGGCTTCTACAGGAAATGTGTATCCACA 3110
Qy 2895 GCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCA 2954
Db 3111 GCCTCTTCTGAACTCTAGCATGTGTTTAGAGGACAGCAGATACAACTGGTCTTTCATTCA 3170
Qy 2955 CTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAA 3014
Db 3171 CTGTCCAGCTTGCCAGTGCAATGGACACAGCAAATGTATTAACCAGAGTATCTGTGAGAA 3230
Qy 3015 GTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGA 3074
Db 3231 GTGTGAGAACCTGACCACTGGCAAGCACTGTGAGACCTGCATATCTGGCTTCTATGGTGA 3290
Qy 3075 TCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAA 3134
Db 3291 TCCGACTAATGGAGGCAAATGTCAGCCATGTAAGTGCAACGGGCATGCATCGCTGTGCAA 3350
Qy 3135 CACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCT 3194
Db 3351 TACCAACACCGGCAAGTGCTTCTGCACCACCAAGGGTGTCAAGGGAGATGAGTGCCAGCT 3410
Qy 3195 ATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCT 3254
Db 3411 ATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAAAGGAACATGTTATTATACCCT 3470
Qy 3255 TCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGC 3314
Db 3471 TCTCATTGACTATCAGTTCACTTTTAGCCTGTCCCAGGAAGACGATCGCTATTATACAGC 3530
Qy 3315 TATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGC 3374
Db 3531 CATCAATTTTGTGGCCACTCCTGATGAACAAAACAGGGATTTGGACATGTTTATCAATGC 3590

```
Qy       3375 CTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGC 3434
Db       3591 CTCAAAAAACTTCAACCTCAACATCACCTGGGCTACCAGCTTCTCAGCTGGAACCCAGGC 3650

Qy       3435 TGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTC 3494
Db       3651 AGGAGAAGAAATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGCTTCTC 3710

Qy       3495 TAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAA 3554
Db       3711 TAATGAGAAATTTGATTTTCGCAACCATCCAAACATTACTTTCTTTGTTTATGTCAGTAA 3770

Qy       3555 TTTCACCTGGCCCATCAAAATTCAGGT 3581
Db       3771 TTTCACCTGGCCCATCAAAATTCAGAT 3797
```

RESULT 8
AF119821
LOCUS AF119821 4313 bp mRNA linear ROD 16-APR-1999
DEFINITION Mus musculus attractin (Mgca) mRNA, complete cds.
ACCESSION AF119821
VERSION AF119821.1 GI:4585306
KEYWORDS .
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 4313)
AUTHORS Gunn,T.M., Miller,K.A., He,L., Hyman,R.W., Davis,R.W., Azarani,A., Schlossman,S.F., Duke-Cohan,J.S. and Barsh,G.S.
TITLE The mouse mahogany locus encodes a transmembrane form of human attractin
JOURNAL Nature 398 (6723), 152-156 (1999)
MEDLINE 99184160
PUBMED 10086356
REFERENCE 2 (bases 1 to 4313)
AUTHORS Gunn,T.M., Miller,K.A., He,L., Hyman,R.W., Davis,R.W., Azarani,A., Schlossman,S.F., Duke-Cohan,J.S. and Barsh,G.S.
TITLE Direct Submission
JOURNAL Submitted (12-JAN-1999) HHMI, Stanford, Beckman Center B271A, Stanford, CA 94305-5323, USA
FEATURES Location/Qualifiers
source 1. .4313
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/chromosome="2"
gene 1. .4313
/gene="Mgca"
/note="mahogany candidate"
CDS 27. .4313
/gene="Mgca"
/note="mahogany"
/codon_start=1
/product="attractin"
/protein_id="AAD25372.1"
/db_xref="GI:4585307"
/translation="MVAVAVAASTEARLRGSTTATAAPAGRKGRQHRPCTATGAWRPG
PRARLCLPRVLSRALPPPPLLPLLFSLLLLPLPREAEAAAVAAAVSGSAAAEAKECDR
PCVNGGRCNPGTGQCVCPTGWVGEQCQHCGGRFRLTGSSGFVTDGPGNYKYKTKCTWL
IEGYPNAVLRLRFNHFATECSWDHLYVYDGDSIYAPLIAAFSGLIVPERDGNETAPEV
TVTSGYALLHFFSDAAYNLTGFNITYNFDMCPNNCSARGECKSSNSSSAVECECSENW
KGESCDIPHCTDNCGFPHRGICNASDTRGCSCFPHWQGPGCSIPVPANQSFWTREEYS
DLKLPRASHKAVVNGNIMWVVGGYMFNHSDYSMVLAYDLTSREWLPLNHSVNSVVVRY
GHSLALHKDKIYMYGGKIDSTGNVTNELRVFHIHNESWVLLTPKAKDQYAVVGHSAHI
VTLASGRVVMLVIFGHCPLYGYISVVQEYDLEKNTWSILHTQGALVQGGYGHSSAYDD
RTKALYVHGGYKAFSANKYRLADDLYRYDVDTQMWTILKDSRFFRYLHTAVIVSGTML
VFGGNTHNDTSMSHGAKCFSSDFMAYDIACDRWSVLPRPELHHDVNRFGHSAVLYNST
MYVFGGFNSLLLSDVLVFTSEQCDAHRSEAACVAAGPGIRCLWDTQSSRCTSWELATE
EQAEKLKSECFSKRTLDHDRCDQHTDCYSCTANTNDCHWCNDHCVPVNHSCTEGQISI
AKYESCPKDNPMYYCNKKTSCRSCALDQNCQWEPRNQECIALPENICGNGWHLVGNSC
LKITTAKENYDNAKLSCRNHNAFLASLTSQKKVELVLKQLRLMQSSQSMSKLTLTPWV
GLRKINVSYWCWEDMSPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLNG
SVCERPANHSAKQCRTPCALRTACGECTSSSSECMWCSNMKQCVDSNAYVASFPFGQC
MEWYTMSSCPPENCSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKMPSQA
SAGNVYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSVCEKCEDLTTGKHC
ETCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGDECQLCKVENRYQ
GNPLKGTCYYTLLIDYQFTFSLSQGDDRYYTAINFVATPDEQNRDFDMFINASKKFNL
NITWATSFPAGTQTGEEVPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTWP
IKIQIAFSQHSNFMDLVQFFVTFFSCFLSLLLVAAVVWKIKQSCWASRRREQLLREMQ
QMASRPFASVNVALETDEEPPDLIGGSIKTVPKPIALEPCFGNKAAVLSVFVRLPRGL
GGIPPPGQSGLAVASALVDISQQMPIVYKEKSGAVRNRKQQPPAQPGTCI"
ORIGIN

Query Match 81.0%; Score 2913.6; DB 10; Length 4313;
Best Local Similarity 88.9%; Pred. No. 0;
Matches 3172; Conservative 0; Mismatches 394; Indels 2; Gaps 2;

```
Qy         15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74
Db        260 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCGGGCTCGGCCGCAGCCGAGGCCAA 319

Qy         75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134
Db        320 GGAATGTGACCGGCCGTGTGTCAACGGCGGCCGCTGCAACCCTGGCACCGGCCAGTGCGT 379

Qy        135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194
Db        380 CTGCCCCACGGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 439

Qy        195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254
Db        440 TGGCTCTTCTGGATTTGTAACAGATGGACCTGGGAATTATAAATATAAGACGAAGTGCAC 499

Qy        255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314
Db        500 ATGGCTCATTGAAGGCTATCCAAATGCAGTGCTAAGGTTAAGATTCAACCATTTTGCTAC 559

Qy        315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374
Db        560 AGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATCTACGCACCTCTGAT 619

Qy        375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434
Db        620 TGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCTCCTGAGGT 679

Qy        435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494
Db        680 CACTGTCACTTCAGGTTATGCACTGCTGCATTTTTTCAGTGATGCTGCTTATAATCTGAC 739

Qy        495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554
Db        740 TGGATTTAATATCACTTACAATTTTGACATGTGTCCGAATAATTGCTCAGCCCGAGGAGA 799

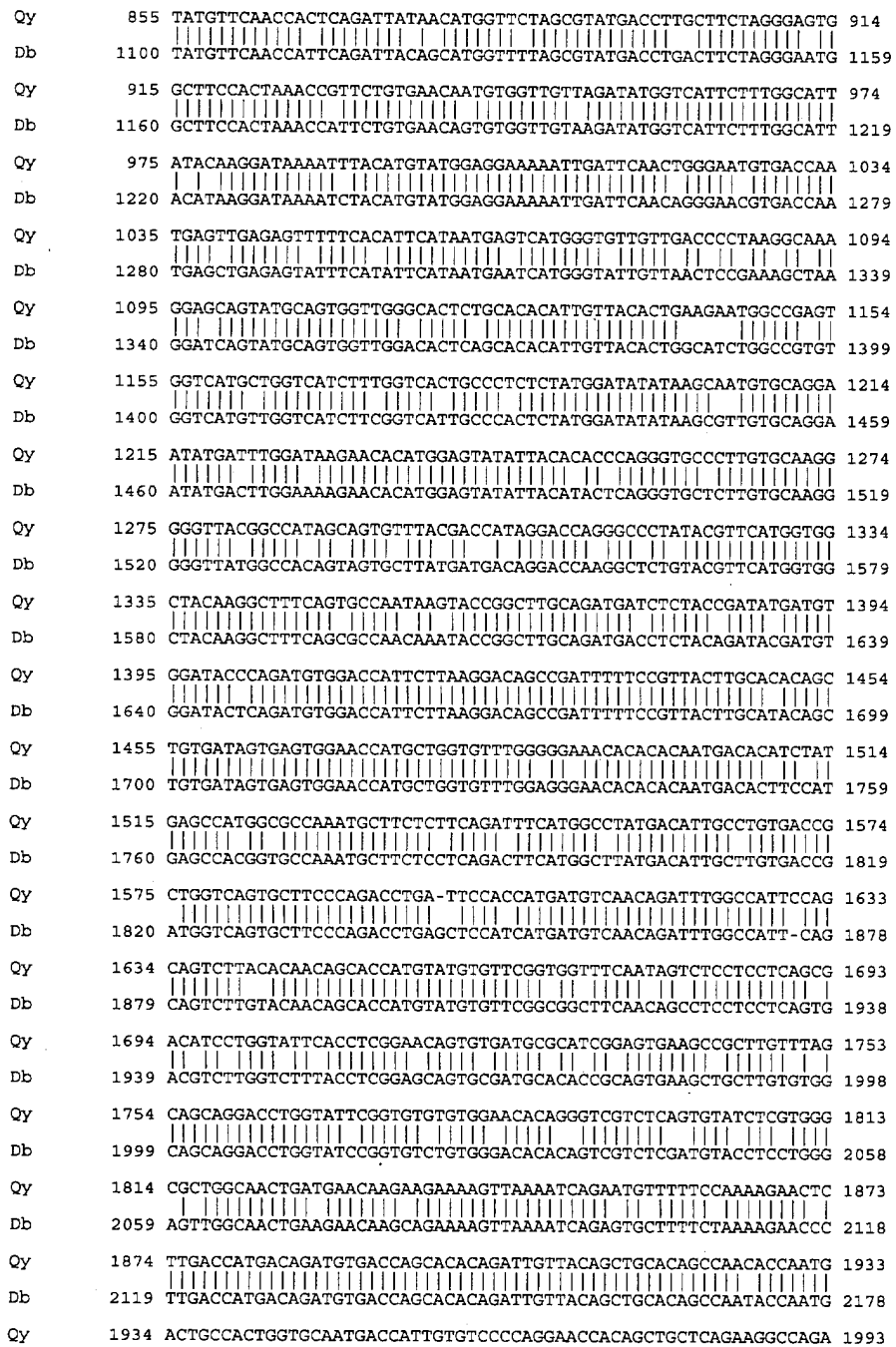
Qy        555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614
Db        800 GTGTAAGAGCAGTAACAGCAGCGCTGTTGAGTGTGAATGTTCTGAAAACTGGAAAGG 859

Qy        615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 674
Db        860 GGAGTCGTGTGACATTCCTCACTGTACAGACAACTGTGGCTTTCCTCACCGAGGCATCTG 919

Qy        675 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 734
Db        920 TAATGCAAGCGATACCAGAGGGTGCTCCTGCTTTCCTCACTGGCAGGGTCCTGGATGCTC 979

Qy        735 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 794
Db        980 AATTCCTGTGCCAGCTAACCAGTCTTTTTGGACTCGAGAAGAATATTCTGATTTAAAGCT 1039

Qy        795 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 854
Db       1040 TCCCAGAGCCTCTCATAAAGCTGTGGTCAATGGAAATATAATGTGGGTTGTTGGCGGATA 1099
```


Qy 855 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 914
Db 1100 TATGTTCAACCATTCAGATTACAGCATGGTTTTAGCGTATGACCTGACTTCTAGGGAATG 1159
Qy 915 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 974
Db 1160 GCTTCCACTAAACCATTCTGTGAACAGTGTGGTTGTAAGATATGGTCATTCTTTGGCATT 1219
Qy 975 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1034
Db 1220 ACATAAGGATAAAATCTACATGTATGGAGGAAAAATTGATTCAACAGGGAACGTGACCAA 1279
Qy 1035 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1094
Db 1280 TGAGCTGAGAGTATTTCATATTCATAATGAATCATGGGTATTGTTAACTCCGAAAGCTAA 1339
Qy 1095 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1154
Db 1340 GGATCAGTATGCAGTGGTTGGACACTCAGCACACATTGTTACACTGGCATCTGGCCGTGT 1399
Qy 1155 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1214
Db 1400 GGTCATGTTGGTCATCTTCGGTCATTGCCCACTCTATGGATATATAAGCGTTGTGCAGGA 1459
Qy 1215 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1274
Db 1460 ATATGACTTGGAAAAGAACACATGGAGTATATTACATACTCAGGGTGCTCTTGTGCAAGG 1519
Qy 1275 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1334
Db 1520 GGGTTATGGCCACAGTAGTGCTTATGATGACAGGACCAAGGCTCTGTACGTTCATGGTGG 1579
Qy 1335 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1394
Db 1580 CTACAAGGCTTTCAGCGCCAACAAATACCGGCTTGCAGATGACCTCTACAGATACGATGT 1639
Qy 1395 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1454
Db 1640 GGATACTCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCATACAGC 1699
Qy 1455 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1514
Db 1700 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGAGGGAACACACACAATGACACTTCCAT 1759
Qy 1515 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1574
Db 1760 GAGCCACGGTGCCAAATGCTTCTCCTCAGACTTCATGGCTTATGACATTGCTTGTGACCG 1819
Qy 1575 CTGGTCAGTGCTTCCCAGACCTGA-TTCCACCATGATGTCAACAGATTTGGCCATTCCAG 1633
Db 1820 ATGGTCAGTGCTTCCCAGACCTGAGCTCCATCATGATGTCAACAGATTTGGCCATT-CAG 1878
Qy 1634 CAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCG 1693
Db 1879 CAGTCTTGTACAACAGCACCATGTATGTGTTCGGCGGCTTCAACAGCCTCCTCCTCAGTG 1938
Qy 1694 ACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAG 1753
Db 1939 ACGTCTTGGTCTTTACCTCGGAGCAGTGCGATGCACACCGCAGTGAAGCTGCTTGTGTGG 1998
Qy 1754 CAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGG 1813
Db 1999 CAGCAGGACCTGGTATCCGGTGTCTGTGGGACACACAGTCGTCTCGATGTACCTCCTGGG 2058
Qy 1814 CGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTC 1873
Db 2059 AGTTGGCAACTGAAGAACAAGCAGAAAAGTTAAAATCAGAGTGCTTTTCTAAAAGAACCC 2118
Qy 1874 TTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATG 1933
Db 2119 TTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAATACCAATG 2178
Qy 1934 ACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGA 1993
Db 2179 ACTGCCACTGGTGCAATGATCACTGTGTCCCTGTGAACCACAGCTGCACAGAAGGCCAGA 2238
Qy 1994 TCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGA 2053
Db 2239 TCTCCATTGCCAAGTATGAGAGTTGCCCCAAGGATAACCCCATGTACTACTGCAATAAGA 2298
Qy 2054 AGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGG 2113
Db 2299 AAACCAGCTGCAGGAGCTGTGCCCTAGACCAGAACTGCCAGTGGGAGCCCCGGAATCAAG 2358
Qy 2114 AGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCAT 2173
Db 2359 AGTGCATCGCCCTGCCGGAAAATATCTGTGGCAATGGCTGGCATTTGGTTGGAAACTCGT 2418
Qy 2174 GTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACC 2233
Db 2419 GTCTGAAAATCACTACTGCTAAGGAGAATTATGACAATGCTAAATTGTCCTGTAGGAACC 2478
Qy 2234 ACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGC 2293
Db 2479 ACAATGCCTTTTTGGCTTCCCTCACATCCCAGAAGAAGGTGGAGCTTGTCCTTAAGCAGC 2538
Qy 2294 TGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCC 2353
Db 2539 TTCGATTAATGCAATCATCTCAAAGTATGTCCAAGCTCACTCTGACTCCATGGGTTGGTC 2598
Qy 2354 TTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTT 2413
Db 2599 TTCGGAAGATCAATGTGTCTTACTGGTGCTGGGAGGATATGTCTCCATTCACAAATAGTT 2658
Qy 2414 TACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAAC 2473
Db 2659 TGCTGCAGTGGATGCCATCTGAGCCCAGTGATGCTGGCTTCTGTGGGATCTTGTCAGAGC 2718
Qy 2474 CCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTG 2533
Db 2719 CTAGTACTCGGGGATTAAAGGCTGCAACCTGCATCAACCCTCTCAATGGCAGCGTCTGTG 2778
Qy 2534 AAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCAT 2593
Db 2779 AAAGGCCTGCAAACCACAGTGCCAAGCAGTGCCGGACACCATGTGCCCTGCGGACAGCGT 2838
Qy 2594 GTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTG 2653
Db 2839 GTGGCGAGTGCACTAGCAGCAGCTCGGAGTGCATGTGGTGCAGTAACATGAAGCAGTGTG 2898
Qy 2654 TGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGA 2713
Db 2899 TGGACTCCAATGCCTACGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGA 2958
Qy 2714 TGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGC 2773
Db 2959 TGAGCAGCTGCCCACCTGAAAATTGCTCTGGCTACTGTACCTGCAGCCATTGCTTGGAGC 3018
Qy 2774 AACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGG 2833
Db 3019 AGCCAGGCTGTGGTTGGTGTACTGATCCTAGCAATACTGGGAAAGGAAAATGTATTGAGG 3078
Qy 2834 GTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCAC 2893
Db 3079 GCAGCTATAAAGGACCTGTGAAGATGCCGTCACAGGCCTCTGCAGGAAATGTGTATCCAC 3138
Qy 2894 AGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTC 2953
Db 3139 AGCCCCTTCTGAACTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTC 3198
Qy 2954 ACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGA 3013
Db 3199 ACTGTCCAGCTTGCCAGTGCAACGGACACAGCAAATGCATCAACCAGAGTGTCTGTGAGA 3258
Qy 3014 AGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTG 3073

```
Db       3259 AGTGTGAGGACCTGACCACGGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTATGGTG 3318

Qy       3074 ATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCA 3133

Db       3319 ACCCGACTAATGGAGGCAAATGTCAGCCATGCAAGTGCAATGGGCACGCATCACTGTGCA 3378

Qy       3134 ACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGC 3193

Db       3379 ACACCAACACCGGCAAGTGCTTCTGTACCACCAAAGGTGTCAAGGGGGACGAGTGCCAGC 3438

Qy       3194 TATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTC 3253

Db       3439 TATGTAAGGTAGAAAATCGATACCAGGGGAACCCTCTCAAAGGAACATGCTACTATACCC 3498

Qy       3254 TTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAG 3313

Db       3499 TTCTCATTGACTATCAGTTCACCTTTAGCCTGTCCCAGGGAGACGACCGCTACTACACAG 3558

Qy       3314 CTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATG 3373

Db       3559 CCATCAACTTTGTGGCTACTCCTGATGAACAAAACAGGGATTTTGACATGTTCATCAATG 3618

Qy       3374 CCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGG 3433

Db       3619 CCTCCAAAAAATTCAACCTCAACATCACCTGGGCCACCAGCTTCCCAGCCGGAACCCAGA 3678

Qy       3434 CTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCT 3493

Db       3679 CTGGAGAAGAGGTGCCTGTTGTTTCAAAAACCAACATCAAGGAATACAAAGATAGCTTCT 3738

Qy       3494 CTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTA 3553

Db       3739 CTAATGAGAAATTTGATTTTCGCAACCATCCAAACATCACTTTCTTTGTTTATGTCAGTA 3798

Qy       3554 ATTTCACCTGGCCCATCAAAATTCAGGT 3581

Db       3799 ATTTCACTTGGCCCATCAAAATTCAGAT 3826
```

```
RESULT 9
AB038387
LOCUS       AB038387               8739 bp    mRNA    linear   ROD 18-JAN-2001
DEFINITION  Rattus norvegicus atrn mRNA for attractin, complete cds.
ACCESSION   AB038387
VERSION     AB038387.1  GI:12275307
KEYWORDS    attractin.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (sites)
  AUTHORS   Kuramoto,T., Kitada,K., Inui,T., Sasaki,Y., Ito,K., Hase,T.,
            Kawaguchi,S., Ogawa,Y., Nakao,K., Barsh,G.S., Nagao,M., Ushijima,T.
            and Serikawa,T.
  TITLE     Attractin/mahogany/zitter plays a critical role in myelination of
            the central nervous system
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (2), 559-564 (2001)
  MEDLINE   21143347
   PUBMED   11209055
REFERENCE   2  (bases 1 to 8739)
  AUTHORS   Kuramoto,T., Serikawa,T. and Ushijima,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-FEB-2000) Takashi Kuramoto, National Cancer Center
            Research Institute, Carcinogenesis Division; Tsukiji 5-1-1,
            Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramot@gan2.ncc.go.jp,
            Tel:81-33542-2511(ex.4521), Fax:81-35565-1753)
FEATURES             Location/Qualifiers
     source          1. .8739
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="WTC"
                     /db_xref="taxon:10116"
     gene            1. .8739
                     /gene="atrn"
     CDS             61. .4359
                     /gene="atrn"
                     /codon_start=1
                     /product="attractin"
                     /protein_id="BAB21017.1"
                     /db_xref="GI:12275308"
                     /translation="MVAAAAAAAEATEARLRGYTTATAAPAGWKERQHRPCAATGAWRP
                     WPRAGLCLPRVLSRALSPPPLLPLLPLLFSLLLLPLPREAEAAAVAAAVSGSAAAEAK
                     ECDRPCVNGGRCNPGTGQCVCPTGWVGEQCQHCGGRFRLTGSSGFVTDGPGNYKYKTK
                     CTWLIEGQPNKIMRLRFNHFATECSWDHLYVYDGDSIYAPLIAAFSGLIVPERDGNET
                     APEVTVTSGYALLHFFSDAAYNLTGFNITYNFDMCPNNCSGRGECKSSNSSSTVECEC
                     SENWKGESCDIPHCTDNCGFPHRGICNASDTRGCSCFPHWQGPGCSIPVPANQSFWTR
                     EEYSDLKLPRASHKAVVNGNIMWVVGGYMFNHSDYSMVLAYDLASREWLSLNHSVNSV
                     VVRYGHSLALHKDKIYMYGGKIDSTGNVTNELRVFHIHNESWVLLTPKAKDQYAVVGH
                     SAHIVTLSSGRVVMLVIFGHCPLYGYISVVQEYDLEKNTWSILQTQGALVQGGYGHSS
                     VYDHRTKALYVHGGYKAFSANKYRLADDLYRYHVDTQMWTILKDSRFFRYLHTAVIVS
                     GTMLVFGGNTHNDTSMSHGAKCFSSDFMAYDIACDRWSVLPRPELHHDVNRFGHSAVL
                     HNSTMYVFGGFNSLLLSDVLVFTSEQCDAHRSEAACVAAGPGIRCLWDTQSSRCTSWE
                     LATEEQAEKLKSECFSKRTLDHDRCDQHTDCYSCTANTNDCHWCNDHCVPVNHSCTEG
                     QISIAKYDNCPKDNPMYYCNKKTSCRSCALDQNCQWEPRNQECIALPENICGIGWHLV
                     GNSCLKITTAKENYDNAKLSCRNHNAFLASLTSQKKVEFVLKQLRLMQSSQSTSKLTL
                     TPWVGLRKINVSYWCWEDMSPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCIN
                     PLNGSVCERPANHSAKQCRTPCALRTACGECTSSSSECMWCSNMKQCVDSNAYVASFP
                     FGQCMEWYTMSSCPPENCSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKM
                     PSHASTGNVYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCEDLTT
                     GKHCETCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGEECQLCEVE
                     NRYQGNPLKGTCYYTLLIDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASK
                     NFNLNITWATSFPAGTQTGEEVPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSN
                     FTWPIKIQIAFSQHSNFMDLVQFFVTFFSCFLSLLLVAAVVWKIKQSCWASRRREQLL
                     REMQQMASRPFASVNVALETDEEPPDLIGGSIKTVPKPIALEPCFGNKAAVLSVFVRL
                     PRGLGGIPPPGQSGLAVASALVDISQQMPIVYKEKSGAVRNRKQQPPAQPGTCI"
ORIGIN
```

```
  Query Match           80.9%;  Score 2909;  DB 10;  Length 8739;
  Best Local Similarity  88.9%;  Pred. No. 0;
  Matches 3168;  Conservative    0;  Mismatches  395;  Indels    2;  Gaps    2;

Qy         15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74

Db        306 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCTGTGTCAGGCTCGGCCGCAGCTGAGGCCAA 365

Qy         75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134

Db        366 GGAATGTGACCGGCCGTGTGTCAACGGCGGCCGCTGCAACCCTGGCACCGGCCAGTGCGT 425

Qy        135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194

Db        426 CTGCCCCACGGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 485

Qy        195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254

Db        486 TGGATCTTCTGGATTTGTAACAGATGGACCTGGGAATTATAAATATAAAACGAAATGCAC 545

Qy        255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314

Db        546 ATGGCTCATTGAAGGACAGCCAAATAAGATAATGAGACTTCGATTCAACCATTTTGCTAC 605

Qy        315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374

Db        606 TGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATTTACGCACCTCTGAT 665

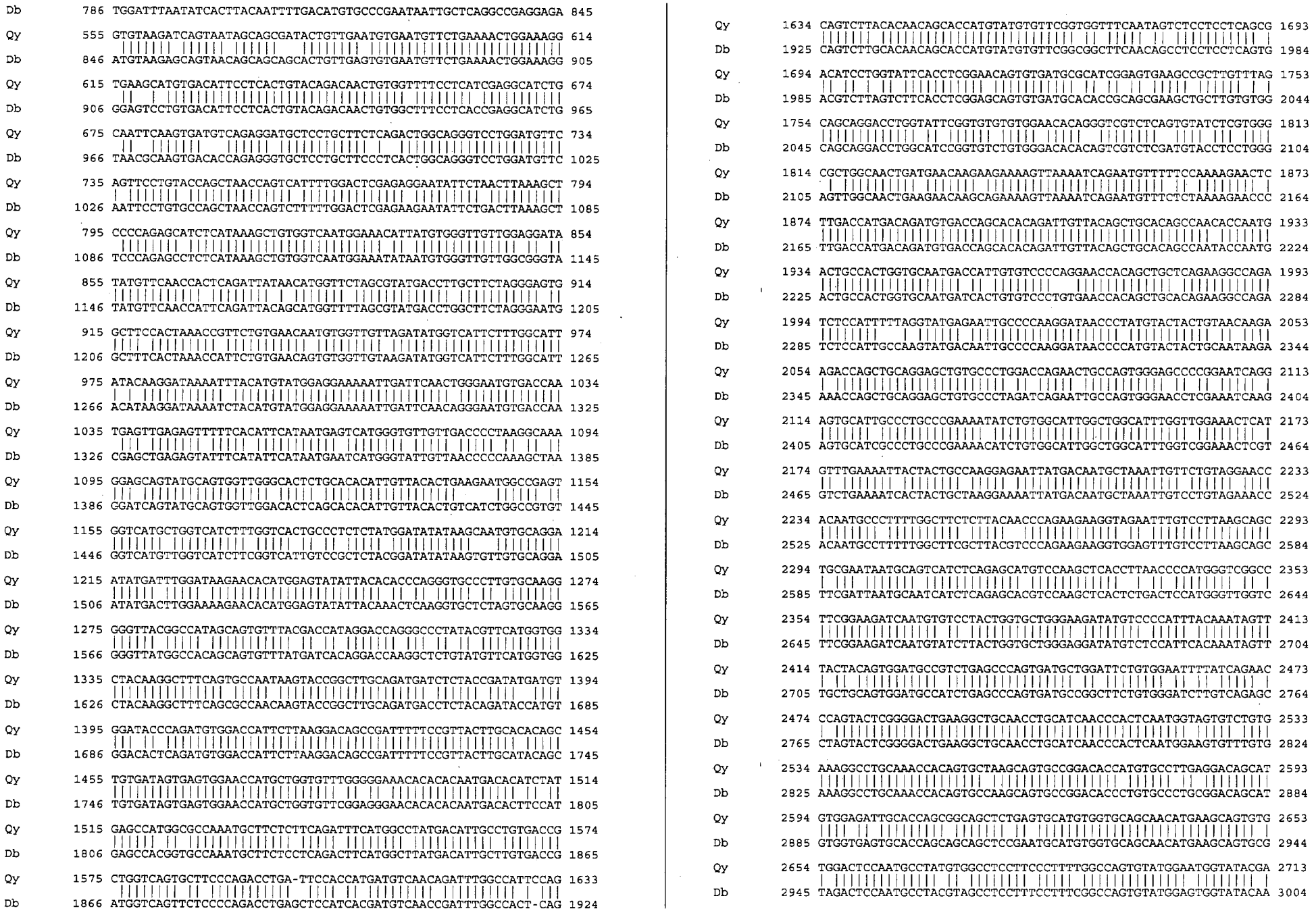
Qy        375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434

Db        666 TGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCCCCTGAAGT 725

Qy        435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494

Db        726 CACTGTCACTTCAGGTTATGCACTACTGCATTTTTTCAGTGATGCTGCTTATAATCTGAC 785

Qy        495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554
```



```
Qy       2714 TGAGCACCTGCCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGC 2773
             |||||| |||||| |||||||||||| || |||||||||||||| || ||||||||||||
Db       3005 TGAGCAGCTGCCCACCTGAAAATTGCTCTGGCTACTGTACCTGCAGCCATTGCTTGGAGC 3064

Qy       2774 AACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGG 2833
             | |||||||||||| |||||||||||||| |||||||||| ||||| ||||| || ||||
Db       3065 AGCCAGGCTGTGGTTGGTGTACTGATCCTAGCAATACTGGGAAAGGCAAATGTATTGAGG 3124

Qy       2834 GTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCAC 2893
             |   |||||||||||| |||||||||||||| || ||| |||||||||| | || ||||
Db       3125 GAAGCTATAAAGGACCCGTGAAGATGCCTTCACATGCCTCTACAGGAAATGTGTACCCAC 3184

Qy       2894 AGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTC 2953
             ||||||| || || |||||||||||||||||||||||||||||||||||||||||||| ||||
Db       3185 AGCCCCTTCTGAACTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTTATTC 3244

Qy       2954 ACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGA 3013
             |||||||||||||||| ||||| || ||||| || |||||||| ||||||||||||||||
Db       3245 ACTGTCCAGCTTGCCAGTGCAATGGACACAGCAAGTGCATCAACCAGAGCATCTGTGAGA 3304

Qy       3014 AGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTG 3073
             |||||||| |||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db       3305 AGTGTGAGGACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTATGGTG 3364

Qy       3074 ATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCA 3133
             | || || |||||||| |||||||||||||||||||||||||||||| |||||||| |||||||
Db       3365 ACCCGACTAATGGAGGCAAATGTCAGCCATGCAAGTGCAACGGGCACGCATCACTGTGCA 3424

Qy       3134 ACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGC 3193
             |||||||||| |||||||||||||| |||||||||||||||||||||| || ||||||||||
Db       3425 ACACCAACACTGGCAAGTGCTTCTGTACCACCAAGGGCGTCAAGGGAGAGGAGTGCCAGC 3484

Qy       3194 TATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCTCAGAGGAACATGTTATTATACTC 3253
             |||||||||||||||||||||||| |||||||||||| |||||||||| |||||| |
Db       3485 TATGTGAGGTAGAAAATCGATACCAGGGAAACCCTCTCAAAGGAACATGTTACTATACCC 3544

Qy       3254 TTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAG 3313
             |||| |||||||||||||||| ||||| || |||||||||| || ||||||||||||||
Db       3545 TTCTCATTGACTATCAGTTCACTTTTAGCCTGTCCCAGGAAGACGACCGCTATTACACAG 3604

Qy       3314 CTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATG 3373
             | |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db       3605 CCATCAATTTTGTGGCTACTCCTGATGAACAAAACAGGGATTTGGACATGTTCATCAATG 3664

Qy       3374 CCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGG 3433
             ||||||| || |||||||||||||||||||||||||  |||||||| ||||||||||||||
Db       3665 CCTCCAAAAACTTCAACCTCAACATCACCTGGGCCACCAGTTTCCCAGCTGGAACCCAGA 3724

Qy       3434 CTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCT 3493
             |||||||||||| |||||||||||||||||||||||||| ||||| |||||||||||| ||||
Db       3725 CTGGAGAAGAGGTGCCTGTTGTTTCAAAAACCAACATCAAGGAATACAAAGATAGCTTCT 3784

Qy       3494 CTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTA 3553
             |||||||||| |||||||||||||||||| ||||| || ||||||||||||||||||||||
Db       3785 CTAATGAGAAATTTGATTTTCGCAACCATCCAAACATTACTTTCTTTGTTTATGTCAGTA 3844

Qy       3554 ATTTCACCTGGCCCATCAAAATTCA 3578
             ||||||| ||||||||||||||||
Db       3845 ATTTCACTTGGCCCATCAAAATTCA 3869
```

RESULT 10
AB038388

```
LOCUS       AB038388               4500 bp    mRNA    linear   ROD 18-JAN-2001
DEFINITION  Rattus norvegicus Atrn mRNA for attractin, complete cds.
ACCESSION   AB038388
VERSION     AB038388.1  GI:12275311
KEYWORDS    attractin.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (sites)
  AUTHORS   Kuramoto,T., Kitada,K., Inui,T., Sasaki,Y., Ito,K., Hase,T.,
            Kawaguchi,S., Ogawa,Y., Nakao,K., Barsh,G.S., Nagao,M., Ushijima,T.
            and Serikawa,T.
  TITLE     Attractin/mahogany/zitter plays a critical role in myelination of
            the central nervous system
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (2), 559-564 (2001)
  MEDLINE   21143347
   PUBMED   11209055
REFERENCE   2  (bases 1 to 4500)
  AUTHORS   Kuramoto,T., Serikawa,T. and Ushijima,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-FEB-2000) Takashi Kuramoto, National Cancer Center
            Research Institute, Carcinogenesis Division; Tsukiji 5-1-1,
            Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramot@gan2.ncc.go.jp,
            Tel:81-33542-2511(ex.4521), Fax:81-35565-1753)
FEATURES             Location/Qualifiers
     source          1. .4500
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="WTC"
                     /db_xref="taxon:10116"
     gene            1. .4500
                     /gene="Atrn"
     CDS             61. .3888
                     /gene="Atrn"
                     /codon_start=1
                     /product="attractin"
                     /protein_id="BAB21018.1"
                     /db_xref="GI:12275312"
                     /translation="MVAAAAAAEATEARLRGYTTATAAPAGWKERQHRPCAATGAWRP
                     WPRAGLCLPRVLSRALSPPPLLPLLPLLFSLLLLPLPREAEAAAVAAAVSGSAAAEAK
                     ECDRPCVNGGRCNPGTGQCVCPTGWVGEQCQHCGGRFRLTGSSGFVTDGPGNYKYKTK
                     CTWLIEGQPNKIMRLRFNHFATECSWDHLYVYDGDSIYAPLIAAFSGLIVPERDGNET
                     APEVTVTSGYALLHFFSDAAYNLTGFNITYNFDMCPNNCSGRGECKSSNSSSTVECEC
                     SENWKGESCDIPHCTDNCGFPHRGICNASDTRGCSCFPHWQGPGCSIPVPANQSFWTR
                     EEYSDLKLPRASHKAEVNGNIMWVVGGYMFNHSDYSMVLAYDLASREWLSLNHSVNSV
                     VVRYGHSLALHKDKIYMYGGKIDSTGNVTNELRVFHIHNESWVLLTPKAKDQYAVVGH
                     SAHIVTLSSGRVVMLVIFGHCPLYGYISVVQEYDLEKNTWSILQTQGALVQGGYGHSS
                     VYDHRTKALYVHGGYKAFSANKYRLADDLYRYHVDTQMWTILKDSRFFRYLHTAVIVS
                     GTMLVFGGNTHNDTSMSHGAKCFSSDFMAYDIACDRWSVLPRPELHHDVNRFGHSAVL
                     HNSTMYVFGGFNSLLLSDVLVFTSEQCDAHRSEAACVAAGPGIRCLWDTQSSRCTSWE
                     LATEEQAEKLKSECFSKRTLDHDRCDQHTDCYSCTANTNDCHWCNDHCVPVNHSCTEG
                     QISIAKYDNCPKDNPMYYCNKKTSCRSCALDQNCQWEPRNQECIALPENICGIGWHLV
                     GNSCLKITTAKENYDNAKLSCRNHNAFLASLTSQKKVEFVLKQLRLMQSSQSTSKLTL
                     TPWVGLRKINVSYWCWEDMSPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCIN
                     PLNGSVCERPANHSAKQCRTPCALRTACGECTSSSSECMWCSNMKQCVDSNAYVASFP
                     FGQCMEWYTMSSCPPENCSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKM
                     PSHASTGNVYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCEDLTT
                     GKHCETCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGEECQLCEVE
                     NRYQGNPLKGTCYYTLLIDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASK
                     NFNLNITWATSFPAGTQTGEEVPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSN
                     FTWPIKIQVRVTS"
ORIGIN

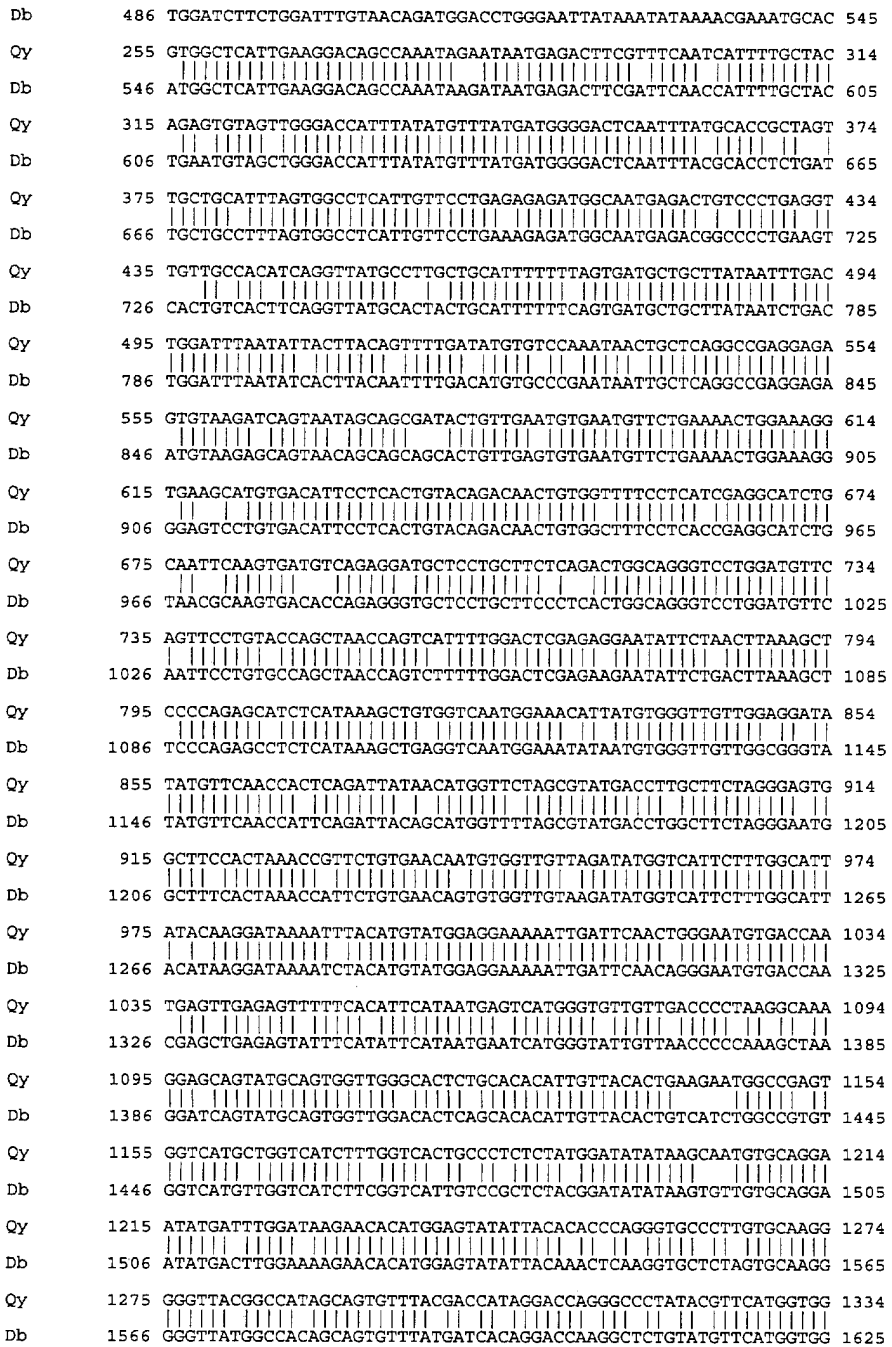
  Query Match             80.8%;  Score 2907.2;  DB 10;  Length 4500;
  Best Local Similarity   88.8%;  Pred. No. 0;
  Matches 3168;  Conservative    0;  Mismatches  398;  Indels    2;  Gaps    2;

Qy         15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74
              || |||     |  ||    | ||  |  |      | | |   |||| ||    | || ||
Db        306 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCTGTGTCAGGCTCGGCCGCAGCTGAGGCCAA 365

Qy         75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134
              |   | |  | ||  |||||||||||||||| |||||||||||||||||||||| ||||||||
Db        366 GGAATGTGACCGGCCGTGTGTCAACGGCGGCCGCTGCAACCCTGGCACCGGGCAGTGCGT 425

Qy        135 CTGCCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194
              ||||||| | |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 CTGCCCCACGGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 485

Qy        195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254
              |||||||||||| ||||| |||||||||||||| |||||||||| |||||||| |||||
```


Db 486 TGGATCTTCTGGATTTGTAACAGATGGACCTGGGAATTATAAATATAAAACGAAATGCAC 545
Qy 255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314
Db 546 ATGGCTCATTGAAGGACAGCCAAATAAGATAATGAGACTTCGATTCAACCATTTTGCTAC 605
Qy 315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374
Db 606 TGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATTTACGCACCTCTGAT 665
Qy 375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434
Db 666 TGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCCCCTGAAGT 725
Qy 435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494
Db 726 CACTGTCACTTCAGGTTATGCACTACTGCATTTTTTCAGTGATGCTGCTTATAATCTGAC 785
Qy 495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554
Db 786 TGGATTTAATATCACTTACAATTTTGACATGTGCCCGAATAATTGCTCAGGCCGAGGAGA 845
Qy 555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614
Db 846 ATGTAAGAGCAGTAACAGCAGCAGCACTGTTGAGTGTGAATGTTCTGAAAACTGGAAAGG 905
Qy 615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 674
Db 906 GGAGTCCTGTGACATTCCTCACTGTACAGACAACTGTGGCTTTCCTCACCGAGGCATCTG 965
Qy 675 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 734
Db 966 TAACGCAAGTGACACCAGAGGGTGCTCCTGCTTCCCTCACTGGCAGGGTCCTGGATGTTC 1025
Qy 735 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 794
Db 1026 AATTCCTGTGCCAGCTAACCAGTCTTTTTGGACTCGAGAAGAATATTCTGACTTAAAGCT 1085
Qy 795 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 854
Db 1086 TCCCAGAGCCTCTCATAAAGCTGAGGTCAATGGAAATATAATGTGGGTTGTTGGCGGGTA 1145
Qy 855 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 914
Db 1146 TATGTTCAACCATTCAGATTACAGCATGGTTTTAGCGTATGACCTGGCTTCTAGGGAATG 1205
Qy 915 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 974
Db 1206 GCTTTCACTAAACCATTCTGTGAACAGTGTGGTTGTAAGATATGGTCATTCTTTGGCATT 1265
Qy 975 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1034
Db 1266 ACATAAGGATAAAATCTACATGTATGGAGGAAAAATTGATTCAACAGGGAATGTGACCAA 1325
Qy 1035 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1094
Db 1326 CGAGCTGAGAGTATTTCATATTCATAATGAATCATGGGTATTGTTAACCCCCAAAGCTAA 1385
Qy 1095 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1154
Db 1386 GGATCAGTATGCAGTGGTTGGACACTCAGCACACATTGTTACACTGTCATCTGGCCGTGT 1445
Qy 1155 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1214
Db 1446 GGTCATGTTGGTCATCTTCGGTCATTGTCCGCTCTACGGATATATAAGTGTTGTGCAGGA 1505
Qy 1215 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1274
Db 1506 ATATGACTTGGAAAAGAACACATGGAGTATATTACAAACTCAAGGTGCTCTAGTGCAAGG 1565
Qy 1275 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1334
Db 1566 GGGTTATGGCCACAGCAGTGTTTATGATCACAGGACCAAGGCTCTGTATGTTCATGGTGG 1625
Qy 1335 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1394
Db 1626 CTACAAGGCTTTCAGCGCCAACAAGTACCGGCTTGCAGATGACCTCTACAGATACCATGT 1685
Qy 1395 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1454
Db 1686 GGACACTCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCATACAGC 1745
Qy 1455 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1514
Db 1746 TGTGATAGTGAGTGGAACCATGCTGGTGTTCGGAGGGAACACACACAATGACACTTCCAT 1805
Qy 1515 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1574
Db 1806 GAGCCACGGTGCCAAATGCTTCTCCTCAGACTTCATGGCTTATGACATTGCTTGTGACCG 1865
Qy 1575 CTGGTCAGTGCTTCCCAGACCTGA-TTCCACCATGATGTCAACAGATTTGGCCATTCCAG 1633
Db 1866 ATGGTCAGTTCTCCCCAGACCTGAGCTCCATCACGATGTCAACCGATTTGGCCACT-CAG 1924
Qy 1634 CAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCG 1693
Db 1925 CAGTCTTGCACAACAGCACCATGTATGTGTTCGGCGGCTTCAACAGCCTCCTCCTCAGTG 1984
Qy 1694 ACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAG 1753
Db 1985 ACGTCTTAGTCTTCACCTCGGAGCAGTGTGATGCACACCGCAGCGAAGCTGCTTGTGTGG 2044
Qy 1754 CAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGG 1813
Db 2045 CAGCAGGACCTGGCATCCGGTGTCTGTGGGACACACAGTCGTCTCGATGTACCTCCTGGG 2104
Qy 1814 CGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTC 1873
Db 2105 AGTTGGCAACTGAAGAACAAGCAGAAAAGTTAAAATCAGAATGTTTCTCTAAAAGAACCC 2164
Qy 1874 TTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATG 1933
Db 2165 TTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAATACCAATG 2224
Qy 1934 ACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGA 1993
Db 2225 ACTGCCACTGGTGCAATGATCACTGTGTCCCTGTGAACCACAGCTGCACAGAAGGCCAGA 2284
Qy 1994 TCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGA 2053
Db 2285 TCTCCATTGCCAAGTATGACAATTGCCCCAAGGATAACCCCATGTACTACTGCAATAAGA 2344
Qy 2054 AGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGG 2113
Db 2345 AAACCAGCTGCAGGAGCTGTGCCCTAGATCAGAATTGCCAGTGGGAACCTCGAAATCAAG 2404
Qy 2114 AGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCAT 2173
Db 2405 AGTGCATCGCCCTGCCCGAAAACATCTGTGGCATTGGCTGGCATTTGGTCGGAAACTCGT 2464
Qy 2174 GTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACC 2233
Db 2465 GTCTGAAAATCACTACTGCTAAGGAAAATTATGACAATGCTAAATTGTCCTGTAGAAACC 2524
Qy 2234 ACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGC 2293
Db 2525 ACAATGCCTTTTTGGCTTCGCTTACGTCCCAGAAGAAGGTGGAGTTTGTCCTTAAGCAGC 2584
Qy 2294 TGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCC 2353
Db 2585 TTCGATTAATGCAATCATCTCAGAGCACGTCCAAGCTCACTCTGACTCCATGGGTTGGTC 2644
Qy 2354 TTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTT 2413
Db 2645 TTCGGAAGATCAATGTATCTTACTGGTGCTGGGAGGATATGTCTCCATTCACAAATAGTT 2704

```
Qy       2414 TACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAAC 2473
Db       2705 TGCTGCAGTGGATGCCATCTGAGCCCAGTGATGCCGGCTTCTGTGGGATCTTGTCAGAGC 2764

Qy       2474 CCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTG 2533
Db       2765 CTAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGAAGTGTTTGTG 2824

Qy       2534 AAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCAT 2593
Db       2825 AAAGGCCTGCAAACCACAGTGCCAAGCAGTGCCGGACACCCTGTGCCCTGCGGACAGCAT 2884

Qy       2594 GTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTG 2653
Db       2885 GTGGTGAGTGCACCAGCAGCAGCTCCGAATGCATGTGGTGCAGCAACATGAAGCAGTGCG 2944

Qy       2654 TGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGA 2713
Db       2945 TAGACTCCAATGCCTACGTAGCCTCCTTTCCTTTCGGCCAGTGTATGGAGTGGTATACAA 3004

Qy       2714 TGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGC 2773
Db       3005 TGAGCAGCTGCCCACCTGAAAATTGCTCTGGCTACTGTACCTGCAGCCATTGCTTGGAGC 3064

Qy       2774 AACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGG 2833
Db       3065 AGCCAGGCTGTGGTTGGTGTACTGATCCTAGCAATACTGGGAAAGGCAAATGTATTGAGG 3124

Qy       2834 GTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCAC 2893
Db       3125 GAAGCTATAAAGGACCCGTGAAGATGCCTTCACATGCCTCTACAGGAAATGTGTACCCAC 3184

Qy       2894 AGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTC 2953
Db       3185 AGCCCCTTCTGAACTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTTATTC 3244

Qy       2954 ACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGA 3013
Db       3245 ACTGTCCAGCTTGCCAGTGCAATGGACACAGCAAGTGCATCAACCAGAGCATCTGTGAGA 3304

Qy       3014 AGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTG 3073
Db       3305 AGTGTGAGGACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTATGGTG 3364

Qy       3074 ATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCA 3133
Db       3365 ACCCGACTAATGGAGGCAAATGTCAGCCATGCAAGTGCAACGGGCACGCATCACTGTGCA 3424

Qy       3134 ACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGC 3193
Db       3425 ACACCAACACTGGCAAGTGCTTCTGTACCACCAAGGGCGTCAAGGGAGAGGAGTGCCAGC 3484

Qy       3194 TATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTC 3253
Db       3485 TATGTGAGGTAGAAAATCGATACCAGGGAAACCCTCTCAAAGGAACATGTTACTATACCC 3544

Qy       3254 TTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAG 3313
Db       3545 TTCTCATTGACTATCAGTTCACTTTTAGCCTGTCCCAGGAAGACGACCGCTATTACACAG 3604

Qy       3314 CTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATG 3373
Db       3605 CCATCAATTTTGTGGCTACTCCTGATGAACAAAACAGGGATTTGGACATGTTCATCAATG 3664

Qy       3374 CCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGG 3433
Db       3665 CCTCCAAAAACTTCAACCTCAACATCACCTGGGCCACCAGTTTCCCAGCTGGAACCCAGA 3724

Qy       3434 CTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCT 3493
Db       3725 CTGGAGAAGAGGTGCCTGTTGTTTCAAAAACCAACATCAAGGAATACAAAGATAGCTTCT 3784

Qy       3494 CTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTA 3553
Db       3785 CTAATGAGAAATTTGATTTTCGCAACCATCCAAACATTACTTTCTTTGTTTATGTCAGTA 3844

Qy       3554 ATTTCACCTGGCCCATCAAAATTCAGGT 3581
Db       3845 ATTTCACTTGGCCCATCAAAATTCAAGT 3872
```

RESULT 11
AR164807
LOCUS AR164807 8827 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 1 from patent US 6274339.
ACCESSION AR164807
VERSION AR164807.1 GI:16238010
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 8827)
AUTHORS Moore,K. and Nagle,D.Lynn.
TITLE Methods and compositions for the diagnosis and treatment of body weight disorders, including obesity
JOURNAL Patent: US 6274339-A 1 14-AUG-2001;
FEATURES Location/Qualifiers
source 1. .8827
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN

Query Match 79.1%; Score 2845.2; DB 6; Length 8827;
Best Local Similarity 90.3%; Pred. No. 0;
Matches 3064; Conservative 0; Mismatches 328; Indels 2; Gaps 2;

```
Qy        189 ACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAA 248
Db        677 ACTAACTGGCTCTTCTGGATTTGTAACAGATGGACCTGGGAATTATAAATATAAGACGAA 736

Qy        249 GTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTT 308
Db        737 GTGCACATGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGCTTCAACCATTT 796

Qy        309 TGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACC 368
Db        797 TGCTACAGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATCTACGCACC 856

Qy        369 GCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCC 428
Db        857 TCTGATTGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCTCC 916

Qy        429 TGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAA 488
Db        917 TGAGGTCACTGTCACTTCAGGTTATGCACTGCTGCATTTTTTCAGTGATGCTGCTTATAA 976

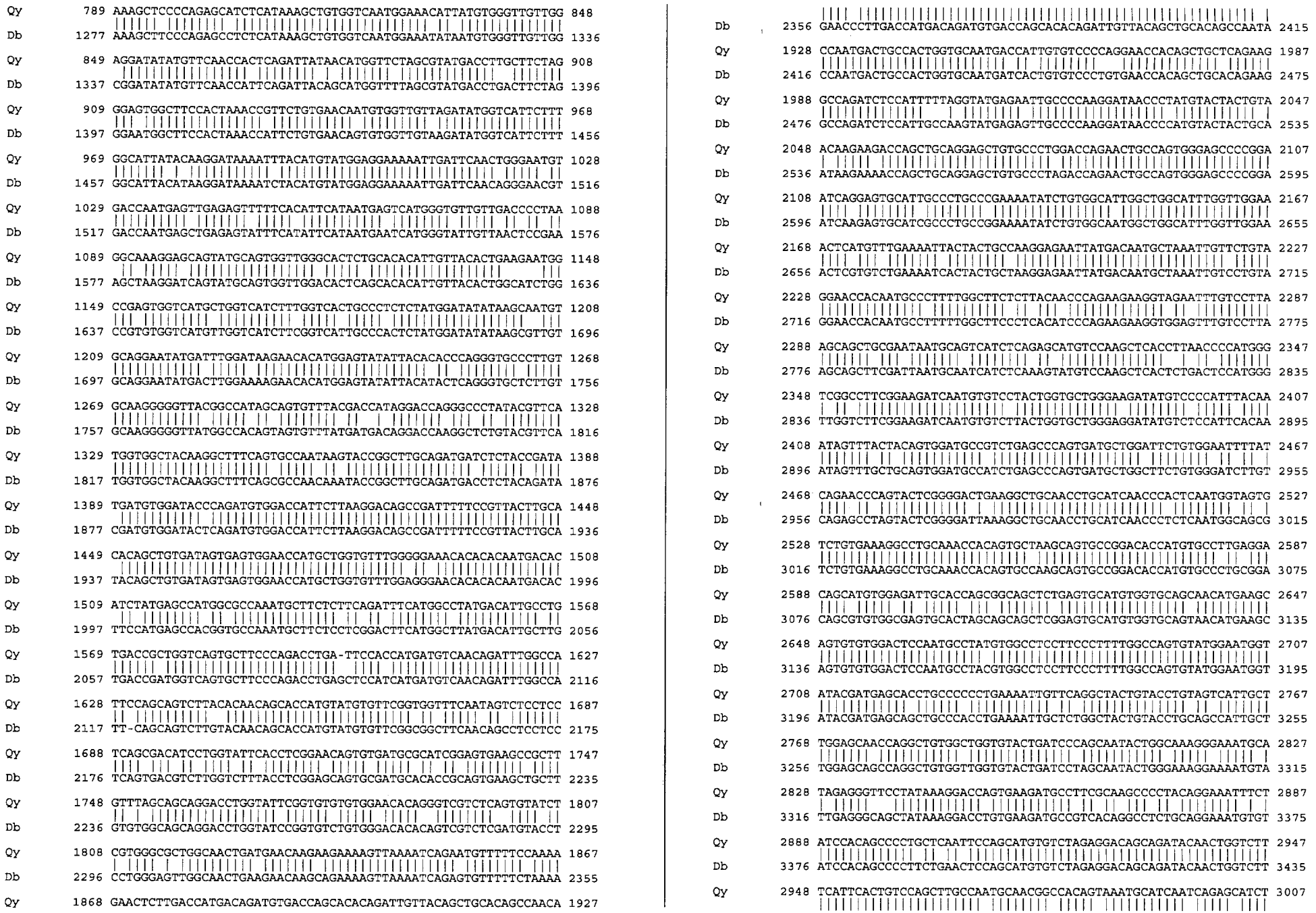
Qy        489 TTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCG 548
Db        977 TCTGACTGGATTTAATATCACTTACAATTTTGACATGTGTCCGAATAATTGCTCAGGCCG 1036

Qy        549 AGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTG 608
Db        1037 AGGAGAGTGTAAGAGCAGTAACAGCAGCAGCGCTGTTGAGTGTGAATGTTCTGAAAACTG 1096

Qy        609 GAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGG 668
Db        1097 GAAAGGGGAGTCGTGTGACATTCCTCACTGTACAGACAACTGTGGCTTTCCTCACCGAGG 1156

Qy        669 CATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGG 728
Db        1157 CATCTGTAATGCAAGCGATACCAGAGGGTGCTCCTGCTTTCCTCACTGGCAGGGTCCTGG 1216

Qy        729 ATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTT 788
Db        1217 ATGTTCAATTCCTGTGCCAGCTAACCAGTCTTTTTGGACTCGAGAAGAATATTCTGATTT 1276
```



```
Db       3436 TCATTCACTGTCCAGCTTGCCAGTGCAACGGACACAGCAAATGCATCAACCAGAGTATCT 3495

Qy       3008 GTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCT 3067
             |||||||||||||| |||||||||| |||||||||||||||||||||||||||||||||||
Db       3496 GTGAGAAGTGTGAGGACCTGACCACGGGCAAGCACTGCGAGACCTGCATATCTGGCTTCT 3555

Qy       3068 ACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTC 3127
             | ||||| || || |||||||| ||||||||||||||||||||||||||||||| || |
Db       3556 ATGGTGACCCGACTAATGGAGGCAAATGTCAGCCATGCAAGTGCAATGGGCACGCATCAC 3615

Qy       3128 TGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGT 3187
             |||||||||||||||| |||||||||| |||||||| || ||||||||||||||||||||
Db       3616 TGTGCAACACCAACACCGGCAAGTGCTTCTGTACCACCAAAGGTGTCAAGGGGGACGAGT 3675

Qy       3188 GCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATT 3247
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||| || |
Db       3676 GCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAAAGGAACATGCTACT 3735

Qy       3248 ATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATT 3307
             |||| ||||| |||||||||||||||||||||||| || |||||||||||| || ||||| |
Db       3736 ATACCCTTCTCATTGACTATCAGTTCACCTTTAGCCTGTCCCAGGAAGACGACCGCTACT 3795

Qy       3308 ACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCA 3367
             ||||||| ||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db       3796 ACACAGCCATCAACTTTGTGGCTACTCCTGATGAACAAAACAGGGATTTGGACATGTTCA 3855

Qy       3368 TCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAA 3427
             |||||||||||| || |||||||||||||||||||||||| |||| ||| |||| ||||
Db       3856 TCAATGCCTCCAAAAACTTCAACCTCAACATCACCTGGGCCACCAGCTTCCCAGCCGGAA 3915

Qy       3428 CCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATA 3487
             ||||| |||||||||||| |||||||||||||||||||||||| ||||| ||||||||||
Db       3916 CCCAGACTGGAGAAGAGGTGCCTGTTGTTTCAAAAACCAACATCAAGGAATACAAAGATA 3975

Qy       3488 GTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATG 3547
             | |||||||||||||| |||||||||||||||| ||||| ||||||||||||||||||||
Db       3976 GCTTCTCTAATGAGAAATTTGATTTTCGCAACCATCCAAACATCACTTTCTTTGTTTATG 4035

Qy       3548 TCAGTAATTTCACCTGGCCCATCAAAATTCAGGT 3581
             ||||||||||||| |||||||||||||||||| |
Db       4036 TCAGTAATTTCACTTGGCCCATCAAAATTCAGAT 4069
```

```
RESULT 12
AB093245
LOCUS       AB093245               5683 bp    mRNA    linear   ROD 07-FEB-2003
DEFINITION  Mus musculus mRNA for mKIAA0548 protein.
ACCESSION   AB093245
VERSION     AB093245.1  GI:26006172
KEYWORDS    FLI_CDNA.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Okazaki,N., Kikuno,R., Ohara,R., Inamoto,S., Hara,Y., Nagase,T.,
            Ohara,O. and Koga,H.
  TITLE     Prediction of the coding sequences of mouse homologues of KIAA
            gene: I. The complete nucleotide sequences of 100 mouse
            KIAA-homologous cDNAs identified by screening of terminal sequences
            of cDNA clones randomly sampled from size-fractionated libraries
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 5683)
  AUTHORS   Okazaki,N., Kikuno,R., Nagase,T., Ohara,O. and Koga,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-OCT-2002) Hisashi Koga, Kazusa DNA Research
            Institute, Laboratory for Genome Informatics; 2-6-7
            Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan
            (E-mail:mouse@kazusa.or.jp, Tel:81-438-52-3919, Fax:81-438-52-3918)
COMMENT     The CREATE program supported by Japan science and technology
            corporation; cDNA full insert sequencing: Kazusa DNA Research
            Institute; cDNA library construction, clone selection and 5'- &
            3'-end one pass sequencing.
```

```
FEATURES             Location/Qualifiers
     source          1. .5683
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="mbg03041"
                     /tissue_type="brain"
                     /dev_stage="adult"
                     /note="vector:modified pBC SK+"
     gene            1. .5683
                     /gene="mKIAA0548"
     CDS             <1. .3181
                     /gene="mKIAA0548"
                     /note="CDS is predicted by in silico analysis. Start codon
                     is not identified."
                     /codon_start=2
                     /evidence=not_experimental
                     /product="mKIAA0548 protein"
                     /protein_id="BAC41429.1"
                     /db_xref="GI:26006173"
                     /translation="AYDLTSREWLPLNHSVNSVVVRYGHSLALHKDKIYMYGGKIDST
                     GNVTNELRVFHIHNESWVLLTPKAKDQYAVVGHSAHIVTLASGRVVMLVIFGHCPLYG
                     YISVVQEYDLEKNTWSILHTQGALVQGGYGHSSVYDDRTKALYVHGGYKAFSANKYRL
                     ADDLYRYDVDTQMWTILKDSRFFRYLHTAVIVSGTMLVFGGNTHNDTSMSHGAKCFSS
                     DFMAYDIACDRWSVLPRPELHHDVNRFGHSAVLYNSTMYVFGGFNSLLLSDVLVFTLE
                     QCDAHRSEAACVAAGPGIRCLWDTQSSRCTSWELATEEQAEKLKSECFSKRTLDHDRC
                     DQHTDCYSCTANTNDCHWCNDHCVPVNHSCTEGQISIAKYESCPKDNPMYYCNKKTSC
                     RSCALDQNCQWEPRNQECIALPENICGNGWHLVGNSCLKITTAKENYDNAKLSCRNHN
                     AFLASLTSQKKVEFVLKQLRLMQSSQSMSKLTLTPWVGLRKINVSYWCWEDMSPFTNS
                     LLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLNGSVCERPANHSAKQCRTPCALR
                     TACGECTSSSSECMWCSNMKQCVDSNAYVASFPFGQCMEWYTMSSCPPENCSGYCTCS
                     HCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKMPSQASAGNVYPQPLLNSSMCLEDSR
                     YNWSFIHCPACQCNGHSKCINQSVCEKCEDLTTGKHCETCISGFYGDPTNGGKCQPCK
                     CNGHASLCNTNTGKCFCTTKGVKGDECQLCEVENRYQGNPLKGTCYYTLLIDYQFTFS
                     LSQEDDRYYTAINFVATPDEQNRDLDMFINASKNFNLNITWATSFPAGTQTGEEVPVV
                     SKTTIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTWPIKIQIAFSQHSNFMDLVQFFV
                     TFFSCFLSLLLVAAVVWKIKQSCWASRRREQLLREMQQMASRPFASVNVALETDEEPP
                     DLIGGSIKTVPKPIALEPCFGNKAAVLSVFVRLPRGLGGIPPPGQSGLAVASALVDIS
                     QQMPIVYKEKSGAVRNRKQQPPAQPGTCI"
ORIGIN
```

```
  Query Match           62.7%;  Score 2255;  DB 10;  Length 5683;
  Best Local Similarity 90.3%;  Pred. No. 0;
  Matches 2433;  Conservative    0;  Mismatches  260;  Indels    2;  Gaps    2;

Qy        888 AGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGT 947
              |||||||||||  |||||||||| |||||||||||||||| |||||||||| ||||||
Db          1 AGCGTATGACCTGACTTCTAGGGAATGGCTTCCACTAAACCATTCTGTGAACAGTGTGGT 60

Qy        948 TGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAA 1007
              ||| |||||||||||||||||||||| | |||||||||| ||||||||||||||||||
Db         61 TGTAAGATATGGTCATTCTTTGGCATTACATAAGGATAAAATCTACATGTATGGAGGAAA 120

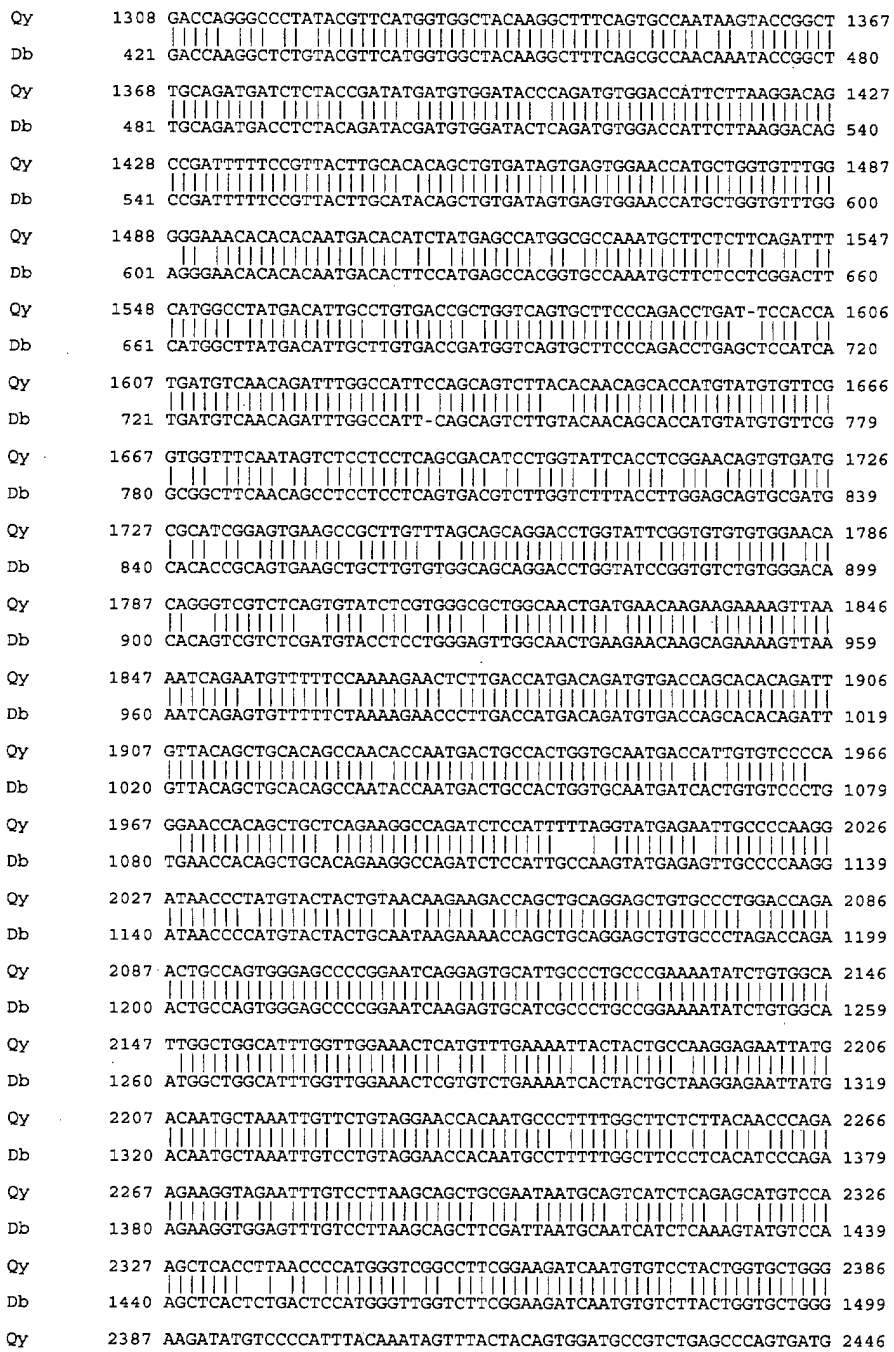
Qy       1008 AATTGATTCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTC 1067
              |||||||||||| ||||| |||||||||||| ||||||| ||||| |||||||||| ||
Db        121 AATTGATTCAACAGGGAACGTGACCAATGAGCTGAGAGTATTTCATATTCATAATGAATC 180

Qy       1068 ATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACA 1127
              |||||| |||| || || || || ||||| |||||||||||||||| ||||| |||||
Db        181 ATGGGTATTGTTAACTCCGAAAGCTAAGGATCAGTATGCAGTGGTTGGACACTCAGCACA 240

Qy       1128 CATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCT 1187
              |||||||||||      |||||| |||||||||| |||||||||| ||||| |||||||
Db        241 CATTGTTACACTGGCATCTGGCCGTGTGGTCATGTTGGTCATCTTCGGTCATTGCCCACT 300

Qy       1188 CTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATT 1247
              |||||||||||||||  |||||||||||||| ||||| ||||||||||||||||||||
Db        301 CTATGGATATATAAGCGTTGTGCAGGAATATGACTTGGAAAAGAACACATGGAGTATATT 360

Qy       1248 ACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAG 1307
              ||| || |||||||| |||||||||||||||| ||||| || |||||||| || | ||
Db        361 ACATACTCAGGGTGCTCTTGTGCAAGGGGGTTATGGCCACAGTAGTGTTTATGATGACAG 420
```


Qy 1308 GACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCT 1367
Db 421 GACCAAGGCTCTGTACGTTCATGGTGGCTACAAGGCTTTCAGCGCCAACAAATACCGGCT 480
Qy 1368 TGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAG 1427
Db 481 TGCAGATGACCTCTACAGATACGATGTGGATACTCAGATGTGGACCATTCTTAAGGACAG 540
Qy 1428 CCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGG 1487
Db 541 CCGATTTTTCCGTTACTTGCATACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGG 600
Qy 1488 GGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTT 1547
Db 601 AGGGAACACACACAATGACACTTCCATGAGCCACGGTGCCAAATGCTTCTCCTCGGACTT 660
Qy 1548 CATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGAT-TCCACCA 1606
Db 661 CATGGCTTATGACATTGCTTGTGACCGATGGTCAGTGCTTCCCAGACCTGAGCTCCATCA 720
Qy 1607 TGATGTCAACAGATTTGGCCATTCCAGCAGTCTTACACAACAGCACCATGTATGTGTTCG 1666
Db 721 TGATGTCAACAGATTTGGCCATT-CAGCAGTCTTGTACAACAGCACCATGTATGTGTTCG 779
Qy 1667 GTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATG 1726
Db 780 GCGGCTTCAACAGCCTCCTCCTCAGTGACGTCTTGGTCTTTACCTTGGAGCAGTGCGATG 839
Qy 1727 CGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACA 1786
Db 840 CACACCGCAGTGAAGCTGCTTGTGTGGCAGCAGGACCTGGTATCCGGTGTCTGTGGGACA 899
Qy 1787 CAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAA 1846
Db 900 CACAGTCGTCTCGATGTACCTCCTGGGAGTTGGCAACTGAAGAACAAGCAGAAAAGTTAA 959
Qy 1847 AATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATT 1906
Db 960 AATCAGAGTGTTTTTCTAAAAGAACCCTTGACCATGACAGATGTGACCAGCACACAGATT 1019
Qy 1907 GTTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCA 1966
Db 1020 GTTACAGCTGCACAGCCAATACCAATGACTGCCACTGGTGCAATGATCACTGTGTCCCTG 1079
Qy 1967 GGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGG 2026
Db 1080 TGAACCACAGCTGCACAGAAGGCCAGATCTCCATTGCCAAGTATGAGAGTTGCCCCAAGG 1139
Qy 2027 ATAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGA 2086
Db 1140 ATAACCCCATGTACTACTGCAATAAGAAAACCAGCTGCAGGAGCTGTGCCCTAGACCAGA 1199
Qy 2087 ACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCA 2146
Db 1200 ACTGCCAGTGGGAGCCCCGGAATCAAGAGTGCATCGCCCTGCCGGAAAATATCTGTGGCA 1259
Qy 2147 TTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATG 2206
Db 1260 ATGGCTGGCATTTGGTTGGAAACTCGTGTCTGAAAATCACTACTGCTAAGGAGAATTATG 1319
Qy 2207 ACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGA 2266
Db 1320 ACAATGCTAAATTGTCCTGTAGGAACCACAATGCCTTTTTGGCTTCCCTCACATCCCAGA 1379
Qy 2267 AGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCA 2326
Db 1380 AGAAGGTGGAGTTTGTCCTTAAGCAGCTTCGATTAATGCAATCATCTCAAAGTATGTCCA 1439
Qy 2327 AGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGG 2386
Db 1440 AGCTCACTCTGACTCCATGGGTTGGTCTTCGGAAGATCAATGTGTCTTACTGGTGCTGGG 1499
Qy 2387 AAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATG 2446
Db 1500 AGGATATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGTGATG 1559
Qy 2447 CTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCA 2506
Db 1560 CTGGCTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGATTAAAGGCTGCAACCTGCA 1619
Qy 2507 TCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCC 2566
Db 1620 TCAACCCTCTCAATGGCAGCGTCTGTGAAAGGCCTGCAAACCACAGTGCCAAGCAGTGCC 1679
Qy 2567 GGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCA 2626
Db 1680 GGACACCATGTGCCCTGCGGACAGCGTGTGGCGAGTGCACTAGCAGCAGCTCGGAGTGCA 1739
Qy 2627 TGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTT 2686
Db 1740 TGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTACGTGGCCTCCTTCCCTT 1799
Qy 2687 TTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCT 2746
Db 1800 TTGGCCAGTGTATGGAATGGTATACGATGAGCAGCTGCCCACCTGAAAATTGCTCTGGCT 1859
Qy 2747 ACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCA 2806
Db 1860 ATTGTACCTGCAGCCATTGCTTGGAGCAGCCAGGCTGTGGTTGGTGTACTGATCCTAGCA 1919
Qy 2807 ATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGC 2866
Db 1920 ATACTGGGAAAGGAAAATGTATTGAGGGCAGCTATAAAGGACCTGTGAAGATGCCGTCAC 1979
Qy 2867 AAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGG 2926
Db 1980 AGGCCTCTGCAGGAAATGTGTATCCACAGCCCCTTCTGAACTCCAGCATGTGTCTAGAGG 2039
Qy 2927 ACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTA 2986
Db 2040 ACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAGTGCAACGGACACAGCA 2099
Qy 2987 AATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCG 3046
Db 2100 AATGCATCAACCAGAGTGTCTGTGAGAAGTGTGAGGACCTGACCACGGGCAAGCACTGCG 2159
Qy 3047 AGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCA 3106
Db 2160 AGACCTGCATATCTGGCTTCTATGGTGACCCGACTAATGGAGGCAAATGTCAGCCATGCA 2219
Qy 3107 AGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCA 3166
Db 2220 AGTGCAATGGGCACGCATCACTGTGCAACACCAACACCGGCAAGTGCTTCTGTACCACCA 2279
Qy 3167 AGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACC 3226
Db 2280 AAGGTGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAGGGAAATC 2339
Qy 3227 CTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTAT 3286
Db 2340 CTCTCAAAGGAACATGCTACTATACCCTTCTCATTGACTATCAGTTCACCTTTAGCCTGT 2399
Qy 3287 CCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAA 3346
Db 2400 CCCAGGAAGACGACCGCTACTACACAGCCATCAACTTTGTGGCTACTCCTGATGAACAAA 2459
Qy 3347 ACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGG 3406
Db 2460 ACAGGGATTTGGACATGTTCATCAATGCCTCCAAAAACTTCAACCTCAACATCACCTGGG 2519
Qy 3407 CTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCA 3466
Db 2520 CCACCAGCTTCCCAGCCGGAACCCAGACTGGAGAAGAGGTGCCTGTTGTTTCAAAAACCA 2579
Qy 3467 ACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAA 3526

```
Db       2580 CCATCAAGGAATACAAAGATAGCTTCTCTAATGAGAAATTTGATTTTCGCAACCATCCAA 2639

Qy       3527 ATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGGT 3581
                |||||||||||||||||||||||||||||||| ||||||||||||||||||| |
Db    .  2640 ACATCACTTTCTTTGTTTATGTCAGTAATTTCACTTGGCCCATCAAAATTCAGAT 2694


RESULT 13
AR164818
LOCUS       AR164818               2625 bp    DNA     linear   PAT 17-OCT-2001
DEFINITION  Sequence 18 from patent US 6274339.
ACCESSION   AR164818
VERSION     AR164818.1  GI:16238029
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2625)
  AUTHORS   Moore,K. and Nagle,D.Lynn.
  TITLE     Methods and compositions for the diagnosis and treatment of body
            weight disorders, including obesity
  JOURNAL   Patent: US 6274339-A 18 14-AUG-2001;
FEATURES             Location/Qualifiers
     source          1. .2625
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             56.4%;  Score 2028.8;  DB 6;  Length 2625;
  Best Local Similarity   97.0%;  Pred. No. 0;
  Matches 2088;  Conservative    0;  Mismatches   62;  Indels    2;  Gaps    2;

Qy         15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74
              || |||    |  ||   | ||  |  |  |  | | |  ||||  ||  || || ||
Db        234 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAA 293

Qy         75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134
              |    | |  | ||  ||||||||||||||||||||||||||||||||||||||||||
Db        294 GGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 353

Qy        135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        354 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 413

Qy        195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 473

Qy        255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 533

Qy        315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 593

Qy        375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 653

Qy        435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        654 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 713

Qy        495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        714 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 773

Qy        555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        774 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 833

Qy        615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        834 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 893

Qy        675 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        894 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 953

Qy        735 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        954 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 1013

Qy        795 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1014 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 1073

Qy        855 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1074 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 1133

Qy        915 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1134 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 1193

Qy        975 ATACAAGGATAAAATTTACATGTATGGAGGAAAAAATTGATTCAACTGGGAATGTGACCAA 1034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1194 ATACAAGGATAAAATTTACATGTATGGAGGAAAAAATTGATTCAACTGGGAATGTGACCAA 1253

Qy       1035 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1254 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1313

Qy       1095 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1314 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1373

Qy       1155 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1374 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1433

Qy       1215 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1434 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1493

Qy       1275 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1494 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1553

Qy       1335 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1554 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1613

Qy       1395 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1454
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1614 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1673

Qy       1455 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1674 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1733

Qy       1515 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1734 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1793

Qy       1575 CTGGTCAGTGCTTCCCAGACCTGAT-TCCACCATGATGTCAACAGATTTGGCCATTCCAG 1633
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||| |||
Db       1794 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATT-CAG 1852

Qy       1634 CAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCG 1693
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1853 CAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCG 1912

Qy       1694 ACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAG 1753
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1913 ACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAG 1972

Qy       1754 CAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGG 1813
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1973 CAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGG 2032

Qy       1814 CGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTC 1873
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2033 CGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTC 2092

Qy       1874 TTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATG 1933
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2093 TTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATG 2152

Qy       1934 ACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGA 1993
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2153 ACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGA 2212

Qy       1994 TCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGA 2053
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2213 TCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGA 2272

Qy       2054 AGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGG 2113
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2273 AGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGG 2332

Qy       2114 AGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGG 2165
              |||||||||||||||||  |  ||  |  |  ||  |  || ||
Db       2333 AGTGCATTGCCCTGCCCGGTAGGCCTTGCAGGGTCATCTTGGTGTGTGTGGG 2384


RESULT 14
AR164813
LOCUS       AR164813               2419 bp    DNA     linear   PAT 17-OCT-2001
DEFINITION  Sequence 8 from patent US 6274339.
ACCESSION   AR164813
VERSION     AR164813.1  GI:16238022
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2419)
  AUTHORS   Moore,K. and Nagle,D.Lynn.
  TITLE     Methods and compositions for the diagnosis and treatment of body
            weight disorders, including obesity
  JOURNAL   Patent: US 6274339-A 8 14-AUG-2001;
FEATURES             Location/Qualifiers
     source          1. .2419
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           39.7%;  Score 1428.8;  DB 6;  Length 2419;
  Best Local Similarity  89.4%;  Pred. No. 0;
  Matches 1560;  Conservative    0;  Mismatches  182;  Indels    2;  Gaps    2;

Qy        189 ACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAA 248
              |||||||| |||||||| ||||| ||||||||||||| |||||||||| || |||||
Db        677 ACTAACTGGCTCTTCTGGATTTGTAACAGATGGACCTGGGAATTATAAATATAAGACGAA 736

Qy        249 GTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTT 308
              |||||| |||||||||||||||||||||||||||||||||||||||||| ||||| |||||
Db        737 GTGCACATGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGCTTCAACCATTT 796

Qy        309 TGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACC 368
              |||||||| |||||||||||||||||||||||||||||||||||||||| || |||||
Db        797 TGCTACAGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATCTACGCACC 856

Qy        369 GCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCC 428
              ||  ||||||| |||||||||||||||||||||| ||||||||||||||||  |  ||
Db        857 TCTGATTGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCTCC 916

Qy        429 TGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAA 488
              ||||||  || ||| |||||||||||  |||||||||||||  |||||||||||||||||
Db        917 TGAGGTCACTGTCACTTCAGGTTATGCACTGCTGCATTTTTTCAGTGATGCTGCTTATAA 976

Qy        489 TTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCG 548
              | |||||||||||||||| |||||| |||||| |||||| ||||||||||||||||||
Db        977 TCTGACTGGATTTAATATCACTTACAATTTTGACATGTGTCCGAATAATTGCTCAGGCCG 1036

Qy        549 AGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTG 608
              |||||||||||||| |||||| ||||||     ||||||| ||||||||||||||||||||
Db       1037 AGGAGAGTGTAAGAGCAGTAACAGCAGCAGCGCTGTTGAGTGTGAATGTTCTGAAAACTG 1096

Qy        609 GAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGG 668
              |||||| ||  | |||||||||||||||||||||||||||||||||| |||||||| ||||
Db       1097 GAAAGGGGAGTCGTGTGACATTCCTCACTGTACAGACAACTGTGGCTTTCCTCACCGAGG 1156

Qy        669 CATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGG 728
              |||||| ||| |||| ||  |||||| |||||||||| | ||||||||||||||||||
Db       1157 CATCTGTAATGCAAGCGATACCAGAGGGTGCTCCTGCTTTCCTCACTGGCAGGGTCCTGG 1216

Qy        729 ATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTT 788
              ||||||| ||||||| |||||||||||||| |||||||||||||| |||||||| | ||
Db       1217 ATGTTCAATTCCTGTGCCAGCTAACCAGTCTTTTTGGACTCGAGAAGAATATTCTGATTT 1276

Qy        789 AAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGG 848
              |||||| |||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       1277 AAAGCTTCCCAGAGCCTCTCATAAAGCTGTGGTCAATGGAAATATAATGTGGGTTGTTGG 1336

Qy        849 AGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAG 908
              |||||||||||||||| |||||||| | ||||||| |||||||||||||  |||||||
Db       1337 CGGATATATGTTCAACCATTCAGATTACAGCATGGTTTTAGCGTATGACCTGACTTCTAG 1396

Qy        909 GGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTT 968
              ||| |||||||||||||||| |||||||||| |||||||| |||||||||||||||||||
Db       1397 GGAATGGCTTCCACTAAACCATTCTGTGAACAGTGTGGTTGTAAGATATGGTCATTCTTT 1456

Qy        969 GGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGT 1028
              ||||||| | |||||||||| |||||||||||||||||||||||||| ||||| ||
Db       1457 GGCATTACATAAGGATAAAATCTACATGTATGGAGGAAAAATTGATTCAACAGGGAACGT 1516

Qy       1029 GACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAA 1088
              ||||||||| |||||| ||||| |||||||||| |||||||| ||||| || || ||
Db       1517 GACCAATGAGCTGAGAGTATTTCATATTCATAATGAATCATGGGTATTGTTAACTCCGAA 1576

Qy       1089 GGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGG 1148
              || |||| |||||||||||||||||||||||||||||||||||||||||||||  |||
Db       1577 AGCTAAGGATCAGTATGCAGTGGTTGGACACTCAGCACACATTGTTACACTGGCATCTGG 1636

Qy       1149 CCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGT 1208
              ||| |||||||| |||||||||| |||| |||| |||||||||||||||||||  |||
Db       1637 CCGTGTGGTCATGTTGGTCATCTTCGGTCATTGCCCACTCTATGGATATATAAGCGTTGT 1696

Qy       1209 GCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGT 1268
              |||||||||| ||||| |||||||||||||||||||| || |||||||| |||||
Db       1697 GCAGGAATATGACTTGGAAAAGAACACATGGAGTATATTACATACTCAGGGTGCTCTTGT 1756

Qy       1269 GCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCA 1328
              |||||||||| ||||| || |||||||| || |||||| ||| || |||||||||||
Db       1757 GCAAGGGGGTTATGGCCACAGTAGTGTTTATGATGACAGGACCAAGGCTCTGTACGTTCA 1816

Qy       1329 TGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATA 1388
              |||||||||||||||||||| ||||| || |||||||||||||||| |||||| ||||
Db       1817 TGGTGGCTACAAGGCTTTCAGCGCCAACAAATACCGGCTTGCAGATGACCTCTACAGATA 1876

Qy       1389 TGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCA 1448
              |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db       1877 CGATGTGGATACTCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCA 1936

Qy       1449 CACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACAC 1508
              |||||||||||||||||||||||||||||||||||| || ||||||||||||||||||
Db       1937 TACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGAGGGAACACACACAATGACAC 1996

Qy       1509 ATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTG 1568
              || |||||||| || |||||||||||||| || || |||||||| |||||||||| ||
```

```
Db       1997 TTCCATGAGCCACGGTGCCAAATGCTTCTCCTCGGACTTCATGGCTTATGACATTGCTTG 2056

Qy       1569 TGACCGCTGGTCAGTGCTTCCCAGACCTGA-TTCCACCATGATGTCAACAGATTTGGCCA 1627
             |||||| |||||||||||||||||||||||   |||| ||||||||||||||||||||||
Db       2057 TGACCGATGGTCAGTGCTTCCCAGACCTGAGCTCCATCATGATGTCAACAGATTTGGCCA 2116

Qy       1628 TTCCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCC 1687
             || |||||||||  ||||||||||||||||||||||||||| || ||||| || |||||||
Db       2117 TT-CAGCAGTCTTGTACAACAGCACCATGTATGTGTTCGGCGGCTTCAACAGCCTCCTCC 2175

Qy       1688 TCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTT 1747
             |||| ||| || |||| || |||||||| ||||| ||||| || || |||||||| ||||
Db       2176 TCAGTGACGTCTTGGTCTTTACCTCGGAGCAGTGCGATGCACACCGCAGTGAAGCTGCTT 2235

Qy       1748 GTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCT 1807
             || | |||||||||||||||||| |||||| ||||| |||||  |||||||  |||| ||
Db       2236 GTGTGGCAGCAGGACCTGGTATCCGGTGTCTGTGGGACACACAGTCGTCTCGATGTACCT 2295

Qy       1808 CGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAA 1867
             | |||| | |||||||||| |||||||  |||||||||||||||||| ||||||| ||||
Db       2296 CCTGGGAGTTGGCAACTGAAGAACAAGCAGAAAAGTTAAAATCAGAGTGTTTTTCTAAAA 2355

Qy       1868 GAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACA 1927
             |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       2356 GAACCCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAATA 2415

Qy       1928 CCAA 1931
              ||||
Db       2416 CCAA 2419


RESULT 15
AR164815
LOCUS       AR164815               6370 bp    DNA     linear   PAT 17-OCT-2001
DEFINITION  Sequence 12 from patent US 6274339.
ACCESSION   AR164815
VERSION     AR164815.1  GI:16238024
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 6370)
  AUTHORS   Moore,K. and Nagle,D.Lynn.
  TITLE     Methods and compositions for the diagnosis and treatment of body
            weight disorders, including obesity
  JOURNAL   Patent: US 6274339-A 12 14-AUG-2001;
FEATURES             Location/Qualifiers
     source          1. .6370
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             37.9%;  Score 1364.4;  DB 6;  Length 6370;
  Best Local Similarity   95.0%;  Pred. No. 0;
  Matches 1492;  Conservative    7;  Mismatches   48;  Indels   23;  Gaps    8;

Qy       2035 ATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAG 2094
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAG 60

Qy       2095 TGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGG 2154
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGG 120

Qy       2155 CATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCT 2214
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCT 180

Qy       2215 AAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTA 2274
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTA 240

Qy       2275 GAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACC 2334
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACC 300

Qy       2335 TTAACCCCATGGGTCGGCCTT---CGGAAGATCAATGTGTCCTACTGG-----TGCTGGG 2386
             |||||||||||||||||| | |    | | | :|||||||||||||||:|     |   |
Db        301 TTAACCCCATGGGTCGGGCCTTCGGGAAGGTYCAATGTGTCCTACTKGGTGCTGGGGAAG 360

Qy       2387 AAGATATGTCCCCATTTACAAATAG-TTTACTACAGTGGATGCCGTCTG----AGCCCAG 2441
             | || :  |||  |||||||||||  |||||||||||||    :  |      ||||||
Db        361 GATATGKTCCCCATTTTACAAATAGTTTTACTACAGTGGGATGSCCGTCTTGAGGCCCAG 420

Qy       2442 TGATGCT--GGATTCTGTGGAATTTT-----ATCAGAACCCAGTACTCGGGGACTGAAGG 2494
             || ||||  |:|||||||| | |||    |  |  |       ||||||||||||||
Db        421 TGTTGCTTGGRATTCTGTGGGAATTTTATTCAGGAACCCAGTTACTTCGGGGACTGAAGG 480

Qy       2495 CTGCAACCTGCAT--CAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAG 2552
             |||||||||||||  | ||   ::|||||||||||||||||||||||||||||||||||||
Db        481 CTGCAACCTGCATTCAACCCACTYMAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAG 540

Qy       2553 TGCTAA-GCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCG 2611
             |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGCTAAGGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCG 600

Qy       2612 GCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATG 2671
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATG 660

Qy       2672 TGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTG 2731
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTG 720

Qy       2732 AAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGT 2791
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGT 780

Qy       2792 GTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAG 2851
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAG 840

Qy       2852 TGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCA 2911
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCA 900

Qy       2912 GCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAAT 2971
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAAT 960

Qy       2972 GCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCA 3031
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCA 1020

Qy       3032 CAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGA 3091
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 CAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGA 1080

Qy       3092 AATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGT 3151
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 AATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGT 1140

Qy       3152 GCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATC 3211
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 GCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATC 1200

Qy       3212 GATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGT 3271
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1201 GATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGT 1260

Qy       3272 TCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTA 3331
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1261 TCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTA 1320

Qy       3332 CTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACC 3391
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1321 CTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACC 1380

Qy       3392 TCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTG 3451
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTG 1440

Qy       3452 TTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATT 3511
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATT 1500

Qy       3512 TTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCA 3571
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCA 1560

Qy       3572 AAATTCAGGT 3581
             |||||||| |
Db       1561 AAATTCAGAT 1570
```

Search completed: March 5, 2004, 05:11:01
Job time : 9065.63 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2004, 16:59:34 ; Search time 5699.24 Seconds
(without alignments)
18847.119 Million cell updates/sec

Title: US-09-787-097-1
Perfect score: 3597
Sequence: 1 atggtggccgcagcggcggc.....aggtgcaaactgaacaatga 3597

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

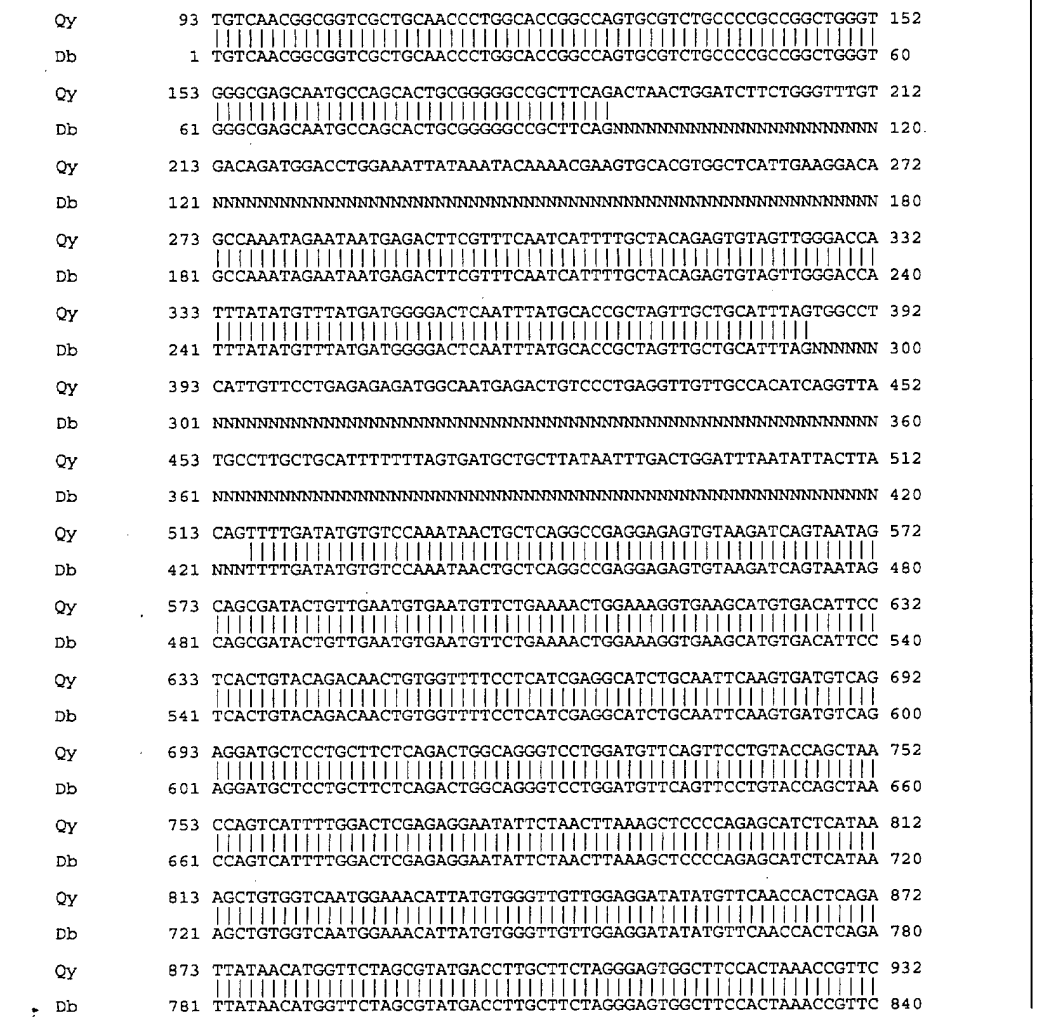
Result No.	Score	% Query Match	Length	DB	ID	Description
1	2942.2	81.8	3976	29	AY418587	AY418587 Homo sapi
2	2508.6	69.7	3671	29	AY418588	AY418588 Pan trogl
3	2507	69.7	3976	29	AY418589	AY418589 Mus muscu
4	925.6	25.7	1201	13	BX440935	BX440935 BX440935
5	775	21.5	964	13	BU506373	BU506373 AGENCOURT
6	763.2	21.2	933	12	BG678679	BG678679 602624477
7	699.6	19.4	840	14	CB520374	CB520374 UI-M-GI0-
8	681	18.9	681	12	BM783739	BM783739 K-EST0061
9	680.4	18.9	683	12	BM783788	BM783788 K-EST0061
10	680	18.9	680	12	BM783258	BM783258 K-EST0061
11	680	18.9	680	12	BM783773	BM783773 K-EST0061
12	677	18.8	677	12	BM783718	BM783718 K-EST0061
13	674.4	18.7	677	12	BM783821	BM783821 K-EST0061
14	673	18.7	673	12	BM783685	BM783685 K-EST0061
15	669.8	18.6	962	13	BX431593	BX431593 BX431593
16	665.4	18.5	792	14	CD644584	CD644584 AGENCOURT
17	636.8	17.7	674	12	BG563729	BG563729 602584512
18	636.4	17.7	770	13	BY761276	BY761276 BY761276
19	632	17.6	632	12	BM773003	BM773003 K-EST0057
20	632	17.6	632	12	BM783264	BM783264 K-EST0061
21	623	17.3	917	10	BF672370	BF672370 602150605
22	618	17.2	633	10	BE894518	BE894518 601433126
23	613	17.0	777	14	CF745736	CF745736 UI-M-GV0-
c 24	612.2	17.0	722	13	BX669806	BX669806 BX669806
25	611	17.0	706	12	BI089332	BI089332 602853292
26	610.4	17.0	953	10	BF313142	BF313142 601897661
27	608	16.9	735	13	BY761390	BY761390 BY761390
28	589.6	16.4	940	13	BQ956223	BQ956223 AGENCOURT
29	585.2	16.3	716	14	CF745901	CF745901 UI-M-GV0-
30	584.8	16.3	991	10	BE880905	BE880905 601490448
c 31	580.2	16.1	714	10	BE650220	BE650220 UI-M-BH2.
32	580	16.1	591	12	BI850289	BI850289 imageqc_1
33	569.8	15.8	653	14	CB446934	CB446934 700912 MA
34	549.6	15.3	902	13	BU151781	BU151781 AGENCOURT
35	547.6	15.2	942	10	BF316716	BF316716 601903293
36	535.2	14.9	595	10	BF853681	BF853681 MR2-EN009
37	534.2	14.9	598	10	BF853006	BF853006 MR2-EN009
38	532.2	14.8	641	10	BB664616	BB664616 BB664616
39	531.2	14.8	663	14	CF745835	CF745835 UI-M-GV0-
40	527.2	14.7	544	10	BF853962	BF853962 MR2-EN009
41	524.2	14.6	626	10	BE980188	BE980188 UI-M-BG2-
42	522.6	14.5	600	14	CB432517	CB432517 608730 MA
c 43	518	14.4	746	14	CB242644	CB242644 UI-CF-FN0
c 44	516	14.3	644	14	CB425680	CB425680 600762 MA
45	508.6	14.1	560	10	BF853965	BF853965 MR2-EN009

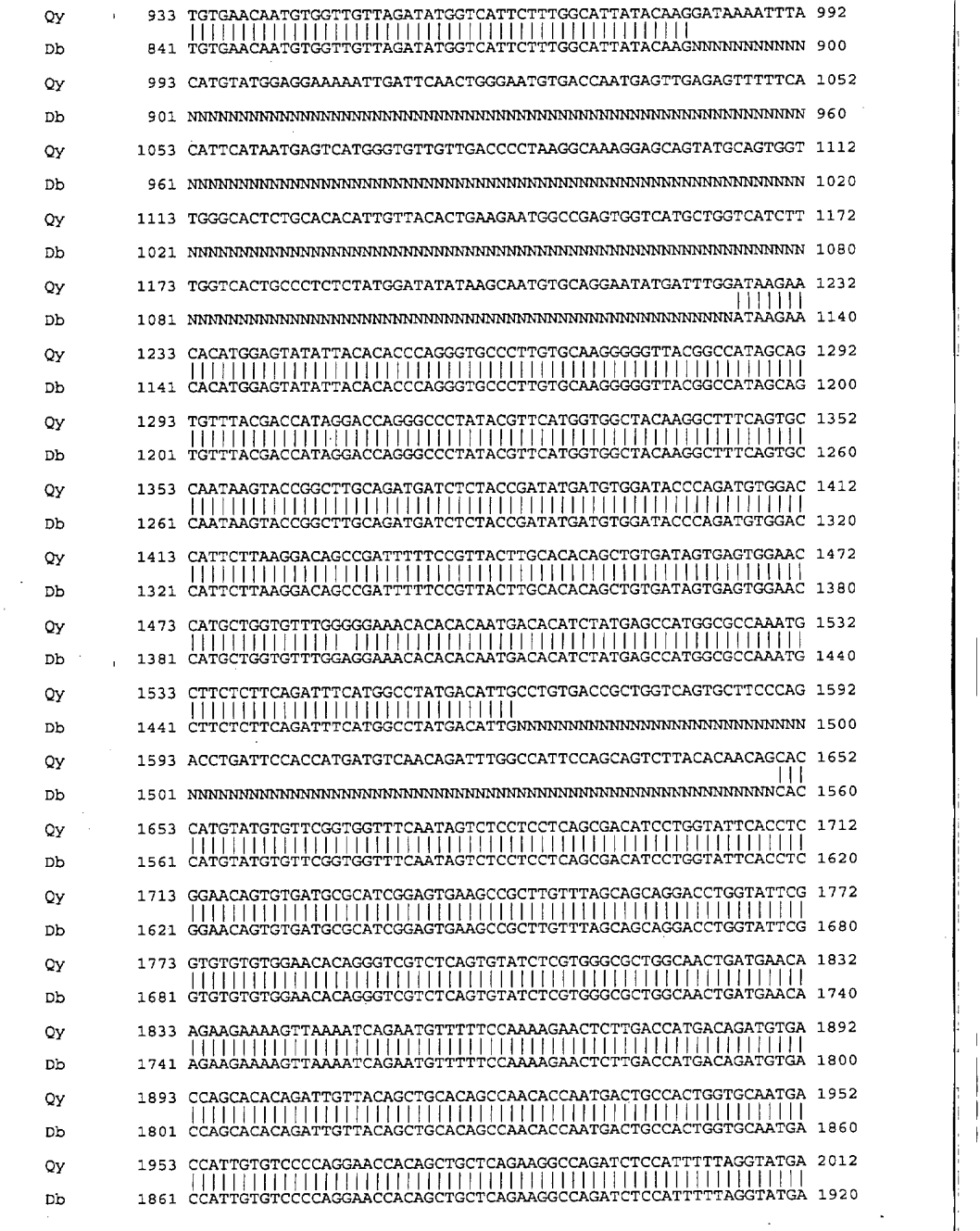
ALIGNMENTS

RESULT 1
AY418587
LOCUS AY418587 3976 bp DNA linear GSS 17-DEC-2003
DEFINITION Homo sapiens ATRN gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION AY418587
VERSION AY418587.1 GI:39774547
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3976)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 3976)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT These sequences were made by sequencing genomic exons and ordering them based on alignment.
FEATURES Location/Qualifiers
source 1. .3976
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
gene <1. .>3976
/gene="ATRN"
/locus_tag="HCM6607"
ORIGIN

Query Match 81.8%; Score 2942.2; DB 29; Length 3976;
Best Local Similarity 84.4%; Pred. No. 0;
Matches 2944; Conservative 0; Mismatches 545; Indels 0; Gaps 0;





```
Qy       2013 GAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTG 2072
             ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db       1921 GAATTGCCCCAAGGATAACCCCATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTG 1980

Qy       2073 TGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGA 2132
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGA 2040

Qy       2133 AAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGC 2192
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 AAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGC 2100

Qy       2193 CAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTC 2252
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 CAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTC 2160

Qy       2253 TCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATC 2312
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 TCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATC 2220

Qy       2313 TCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTC 2372
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 TCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTC 2280

Qy       2373 CTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTC 2432
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 CTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTC 2340

Qy       2433 TGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAA 2492
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 TGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAA 2400

Qy       2493 GGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAG 2552
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 GGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAG 2460

Qy       2553 TGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGG 2612
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 TGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGG 2520

Qy       2613 CAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGT 2672
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 CAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGT 2580

Qy       2673 GGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGA 2732
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 GGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGA 2640

Qy       2733 AAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTG 2792
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 AAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTG 2700

Qy       2793 TACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGT 2852
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 TACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGT 2760

Qy       2853 GAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAG 2912
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 GAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAG 2820

Qy       2913 CATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATG 2972
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 CATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATG 2880

Qy       2973 CAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCAC 3032
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 CAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCAC 2940

Qy       3033 AGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAA 3092
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 AGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAA 3000

Qy       3093 ATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTG 3152
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 ATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTG 3060

Qy       3153 CTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCG 3212
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 CTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCG 3120

Qy       3213 ATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTT 3272
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 ATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTT 3180

Qy       3273 CACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTAC 3332
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 CACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTAC 3240

Qy       3333 TCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCT 3392
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 TCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCT 3300

Qy       3393 CAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGT 3452
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 CAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGT 3360

Qy       3453 TGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTT 3512
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 TGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTT 3420

Qy       3513 TCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAA 3572
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 TCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAA 3480

Qy       3573 AATTCAGGT 3581
             ||||||| |
Db       3481 AATTCAGAT 3489
```

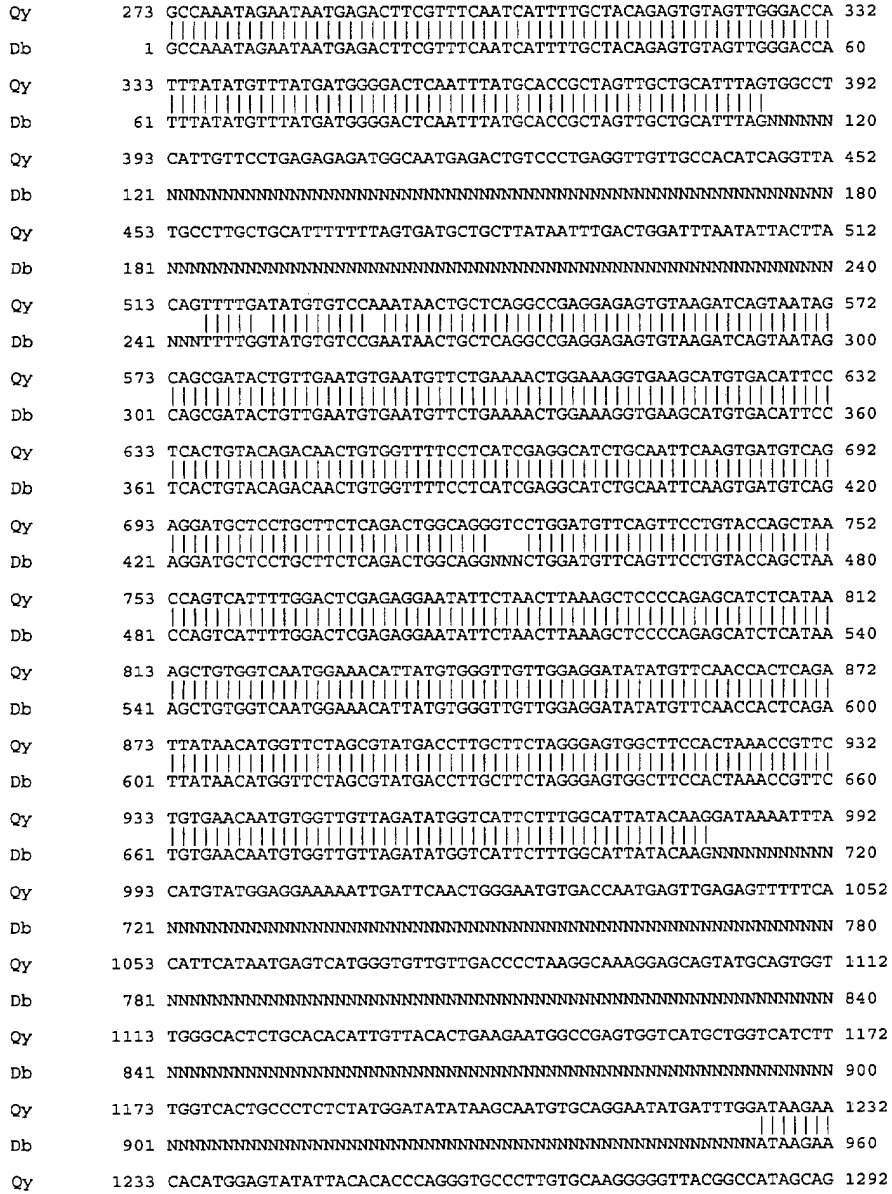
RESULT 2
AY418588

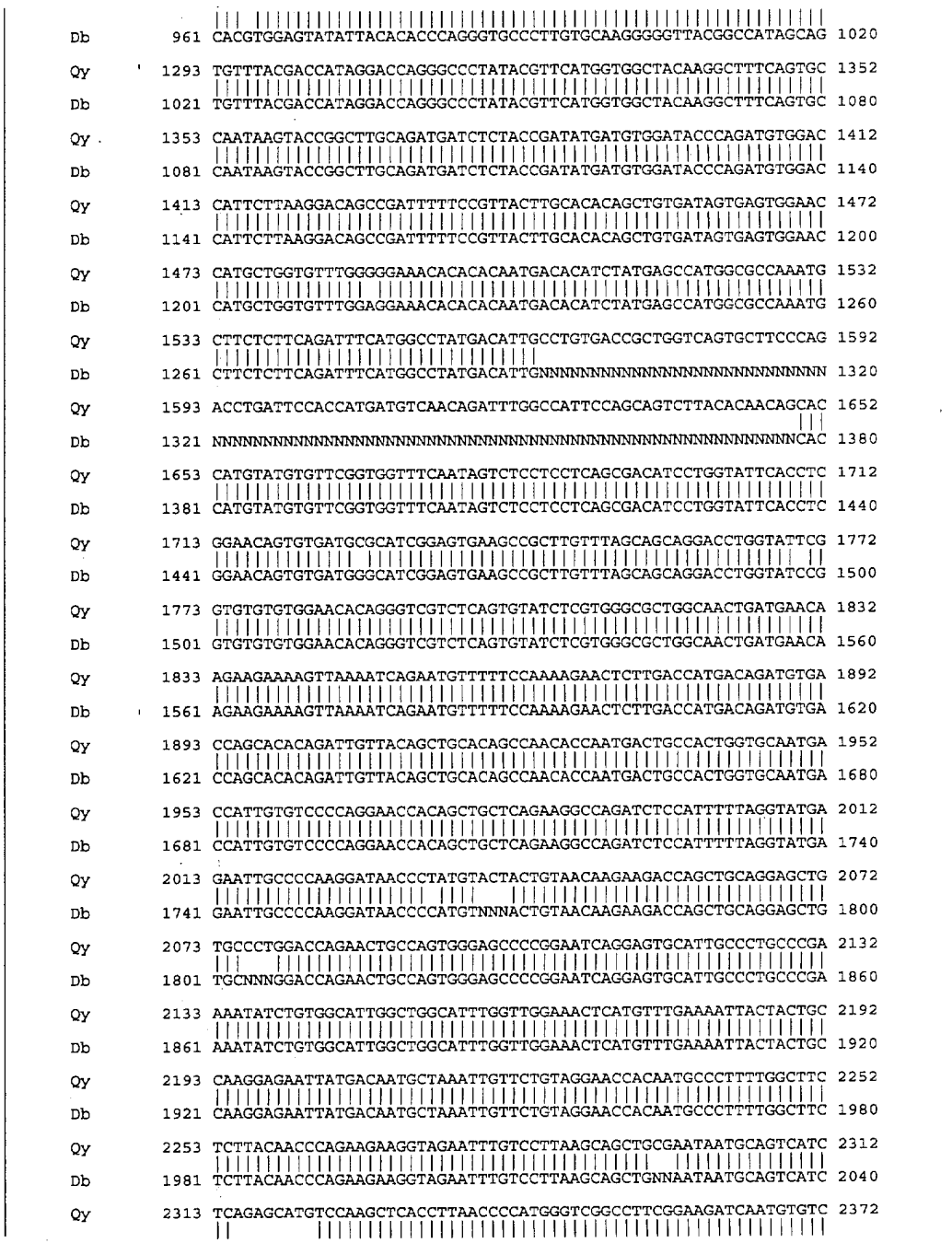
```
LOCUS       AY418588               3671 bp    DNA     linear   GSS 17-DEC-2003
DEFINITION  Pan troglodytes ATRN gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY418588
VERSION     AY418588.1  GI:39774548
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 3671)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 3671)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     These sequences were made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .3671
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
     gene            <1. .>3671
```

/gene="ATRN"
/locus_tag="HCM6607"

ORIGIN

Query Match 69.7%; Score 2508.6; DB 29; Length 3671;
Best Local Similarity 76.1%; Pred. No. 0;
Matches 2517; Conservative 0; Mismatches 792; Indels 0; Gaps 0;





```
Db       2041 TCNNNNNNNNTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTC 2100

Qy       2373 CTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTC 2432
             ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db       2101 CTACTGGTGCTGGGAAGATATGTCTCCATTTACAAATAGTTTACTACAGTGGATGCCGTC 2160

Qy       2433 TGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAA 2492
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 TGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAA 2220

Qy       2493 GGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAG 2552
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       2221 GGCTGCAACCTGCATCAACCCACTCAATGGTAGCGTCTGTGAAAGGCCTGCAAACCACAG 2280

Qy       2553 TGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGG 2612
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 TGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGG 2340

Qy       2613 CAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGT 2672
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       2341 CAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTACGT 2400

Qy       2673 GGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGA 2732
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 GGCCTCCTTCCCTTTCGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCNNNN 2460

Qy       2733 AAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTG 2792

Db       2461 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2520

Qy       2793 TACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGT 2852

Db       2521 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2580

Qy       2853 GAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAG 2912

Db       2581 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2640

Qy       2913 CATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATG 2972
                                                                ||||||||||
Db       2641 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCTTGCCAATG 2700

Qy       2973 CAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCAC 3032
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 CAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCAC 2760

Qy       3033 AGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAA 3092
             |||||||||||||||||||||||||||||||||||| ||||||||||||
Db       2761 AGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTATGGTGATCCCACCANNNNNNNNNN 2820

Qy       3093 ATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTG 3152
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 NNNNNNNNCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTG 2880

Qy       3153 CTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCG 3212
             |||||||||||||||||||||                 ||||||||||||||||||||
Db       2881 CTTCTGCACCACCAAGGGCGTCANNNNNNNNNNNNNNNNNNNNATGTGAGGTAGAAAATCG 2940

Qy       3213 ATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTT 3272
             |||||                 ||||||||||||||||||||||||||
Db       2941 ATACCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNATACTCTTCTTATTGACTATCAGTT 3000

Qy       3273 CACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTAC 3332
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db       3001 CACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATTAATTTTGTGGCTAC 3060

Qy       3333 TCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCT 3392
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 TCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCT 3120

Qy       3393 CAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGT 3452
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 CAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGT 3180

Qy       3453 TGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTT 3512
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 TGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTT 3240

Qy       3513 TCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAA 3572
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 TCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAA 3300

Qy       3573 AATTCAGGT 3581
             ||||||| |
Db       3301 AATTCAGAT 3309
```

```
RESULT 3
AY418589
LOCUS       AY418589                3976 bp    DNA     linear   GSS 17-DEC-2003
DEFINITION  Mus musculus ATRN gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY418589
VERSION     AY418589.1  GI:39774549
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 3976)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 3976)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     These sequences were made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .3976
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
     gene            <1. .>3976
                     /gene="ATRN"
                     /locus_tag="HCM6607"
ORIGIN
```

```
  Query Match           69.7%;  Score 2507;  DB 29;  Length 3976;
  Best Local Similarity   76.6%;  Pred. No. 0;
  Matches 2672;  Conservative    0;  Mismatches  817;  Indels    0;  Gaps    0;

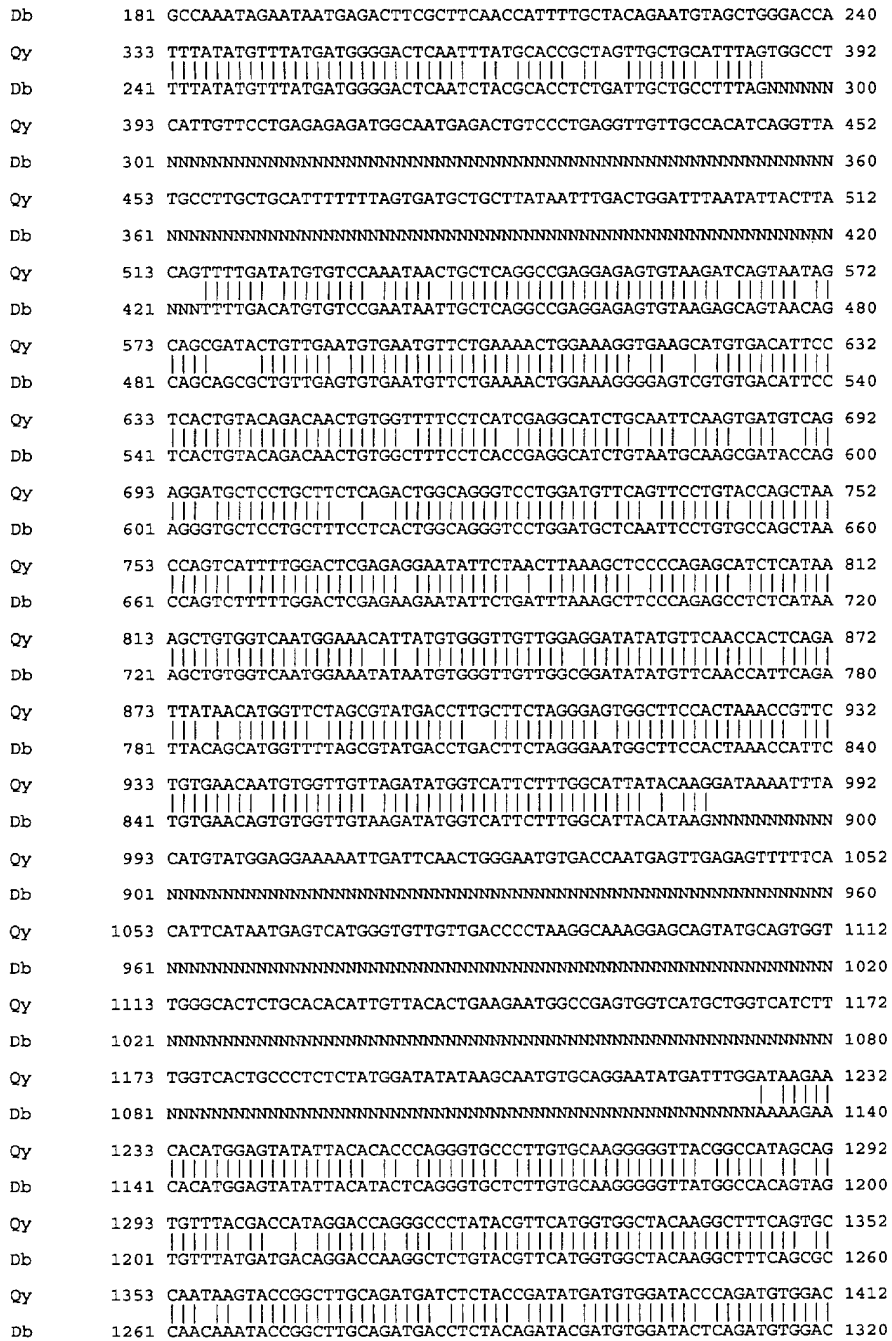
Qy         93 TGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGT 152
              |||||||||||| |||||||||||||||||||||||||||||||||| | ||||||||
Db          1 TGTCAACGGCGGCCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCACGGGCTGGGT 60

Qy        153 GGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGT 212
              ||||||||||||||||||||||||||||||||||||||||
Db         61 GGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGNNNNNNNNNNNNNNNNNNNNNNN 120

Qy        213 GACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGACA 272

Db        121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 180

Qy        273 GCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCA 332
              |||||||||||||||||||||||| ||||| |||||||||||||| ||||| ||||||||
```


Db 181 GCCAAATAGAATAATGAGACTTCGCTTCAACCATTTTGCTACAGAATGTAGCTGGGACCA 240
Qy 333 TTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCT 392
Db 241 TTTATATGTTTATGATGGGGACTCAATCTACGCACCTCTGATTGCTGCCTTTAGNNNNNN 300
Qy 393 CATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTA 452
Db 301 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 360
Qy 453 TGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTA 512
Db 361 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 420
Qy 513 CAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAG 572
Db 421 NNNTTTTGACATGTGTCCGAATAATTGCTCAGGCCGAGGAGAGTGTAAGAGCAGTAACAG 480
Qy 573 CAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCC 632
Db 481 CAGCAGCGCTGTTGAGTGTGAATGTTCTGAAAACTGGAAAGGGGAGTCGTGTGACATTCC 540
Qy 633 TCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAG 692
Db 541 TCACTGTACAGACAACTGTGGCTTTCCTCACCGAGGCATCTGTAATGCAAGCGATACCAG 600
Qy 693 AGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAA 752
Db 601 AGGGTGCTCCTGCTTTCCTCACTGGCAGGGTCCTGGATGCTCAATTCCTGTGCCAGCTAA 660
Qy 753 CCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAA 812
Db 661 CCAGTCTTTTTGGACTCGAGAAGAATATTCTGATTTAAAGCTTCCCAGAGCCTCTCATAA 720
Qy 813 AGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGA 872
Db 721 AGCTGTGGTCAATGGAAATATAATGTGGGTTGTTGGCGGATATATGTTCAACCATTCAGA 780
Qy 873 TTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTC 932
Db 781 TTACAGCATGGTTTAGCGTATGACCTGACTTCTAGGGAATGGCTTCCACTAAACCATTC 840
Qy 933 TGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTA 992
Db 841 TGTGAACAGTGTGGTTGTAAGATATGGTCATTCTTTGGCATTACATAAGNNNNNNNNNN 900
Qy 993 CATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCA 1052
Db 901 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 960
Qy 1053 CATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGT 1112
Db 961 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1020
Qy 1113 TGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTT 1172
Db 1021 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1080
Qy 1173 TGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAA 1232
Db 1081 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNAAAAGAA 1140
Qy 1233 CACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAG 1292
Db 1141 CACATGGAGTATATTACATACTCAGGGTGCTCTTGTGCAAGGGGGTTATGGCCACAGTAG 1200
Qy 1293 TGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGC 1352
Db 1201 TGTTTATGATGACAGGACCAAGGCTCTGTACGTTCATGGTGGCTACAAGGCTTTCAGCGC 1260
Qy 1353 CAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGAC 1412
Db 1261 CAACAAATACCGGCTTGCAGATGACCTCTACAGATACGATGTGGATACTCAGATGTGGAC 1320
Qy 1413 CATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAAC 1472
Db 1321 CATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCATACAGCTGTGATAGTGAGTGGAAC 1380
Qy 1473 CATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATG 1532
Db 1381 CATGCTGGTGTTTGGAGGGAACACACACAATGACACTTCCATGAGCCACGGTGCCAAATG 1440
Qy 1533 CTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAG 1592
Db 1441 CTTCTCCTCAGACTTCATGGCTTATGACATTGNNNNNNNNNNNNNNNNNNNNNNNNNNN 1500
Qy 1593 ACCTGATTCCACCATGATGTCAACAGATTTGGCCATTCCAGCAGTCTTACACAACAGCAC 1652
Db 1501 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCAC 1560
Qy 1653 CATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTC 1712
Db 1561 CATGTATGTGTTCGGCGGCTTCAACAGCCTCCTCCTCAGTGACGTCTTGGTCTTTACCTC 1620
Qy 1713 GGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCG 1772
Db 1621 GGAGCAGTGCGATGCACACCGCAGTGAAGCTGCTTGTGTGGCAGCAGGACCTGGTATCCG 1680
Qy 1773 GTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACA 1832
Db 1681 GTGTCTGTGGGACACACAGTCGTCTCGATGTACCTCCTGGGAGTTGGCAACTGAAGAACA 1740
Qy 1833 AGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGA 1892
Db 1741 AGCAGAAAAGTTAAAATCAGAGTGTTTTTCTAAAAGAACCCTTGACCATGACAGATGTGA 1800
Qy 1893 CCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGCAATGA 1952
Db 1801 CCAGCACACAGATTGTTACAGCTGCACAGCCAATACCAATGACTGCCACTGGTGCAATGA 1860
Qy 1953 CCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGA 2012
Db 1861 TCACTGTGTCCCTGTGAACCACAGCTGCACAGAAGGCCAGATCTCCATTGCCAAGTATGA 1920
Qy 2013 GAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTG 2072
Db 1921 GAGTTGCCCCAAGGATAACCCCATGTACTACTGCAATAAGAAAACCAGCTGCAGGAGCTG 1980
Qy 2073 TGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGA 2132
Db 1981 TGCCCTAGACCAGAACTGCCAGTGGGAGCCCCGGAATCAAGAGTGCATCGCCCTGCCGGA 2040
Qy 2133 AAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGC 2192
Db 2041 AAATATCTGTGGCAATGGCTGGCATTTGGTTGGAAACTCGTGTCTGAAAATCACTACTGC 2100
Qy 2193 CAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTC 2252
Db 2101 TAAGGAGAATTATGACAATGCTAAATTGTCCTGTAGGAACCACAATGCCTTTTTGGCTTC 2160
Qy 2253 TCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATC 2312
Db 2161 CCTCACATCCCAGAAGAAGGTGGAGTTTGTCCTTAAGCAGCTTCGATTAATGCAATCATC 2220
Qy 2313 TCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTC 2372
Db 2221 TCAAAGTATGTCCAAGCTCACTCTGACTCCATGGGTTGGTCTTCGGAAGATCAATGTGTC 2280
Qy 2373 CTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTC 2432
Db 2281 TTACTGGTGCTGGGAGGATATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATC 2340
Qy 2433 TGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAA 2492
Db 2341 TGAGCCCAGTGATGCTGGCTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGATTAAA 2400

```
Qy       2493 GGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAG 2552
             |||||||||||||||||||||| |||||||| || ||||||||||||||||||||||||||
Db       2401 GGCTGCAACCTGCATCAACCCTCTCAATGGCAGCGTCTGTGAAAGGCCTGCAAACCACAG 2460

Qy       2553 TGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGG 2612
             ||| |||||||||||||||||||||||| || ||||||| ||||| || ||||| ||| |
Db       2461 TGCCAAGCAGTGCCGGACACCATGTGCCCTGCGGACAGCGTGTGGCGAGTGCACTAGCAG 2520

Qy       2613 CAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGT 2672
             |||||| |||||||||||||||| |||||||||||||||||||||||||||||||||| ||
Db       2521 CAGCTCGGAGTGCATGTGGTGCAGTAACATGAAGCAGTGTGTGGACTCCAATGCCTACGT 2580

Qy       2673 GGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGA 2732
             |||||||||||||||||||||||||||||||||||||||||||||||||| |||||| ||||
Db       2581 GGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCAGCTGCCCACCTGA 2640

Qy       2733 AAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTG 2792
             |||||| || |||||||||||||| || |||||||||||||| |||||||||||| |||||
Db       2641 AAATTGCTCTGGCTACTGTACCTGCAGCCATTGCTTGGAGCAGCCAGGCTGTGGTTGGTG 2700

Qy       2793 TACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGT 2852
             |||||||| |||||||||| ||||| ||||| || ||||   |||||||||||| ||
Db       2701 TACTGATCCTAGCAATACTGGGAAAGGAAAATGTATTGAGGGCAGCTATAAAGGACCTGT 2760

Qy       2853 GAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAG 2912
             ||||||||| || || ||| || |||||||| | |||||||||||||| || || |||||
Db       2761 GAAGATGCCGTCACAGGCCTCTGCAGGAAATGTGTATCCACAGCCCCTTCTGAACTCCAG 2820

Qy       2913 CATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATG 2972
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       2821 CATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAGTG 2880

Qy       2973 CAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCAC 3032
             |||||| ||||| |||||||||| ||||| |||||||||||||||| |||||||||
Db       2881 CAACGGACACAGCAAATGCATCAACCAGAGTATCTGTGAGAAGTGTGAGGACCTGACCAC 2940

Qy       3033 AGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAA 3092
              |||||||||||||||||||||||||||||||||| ||||| || || |||||||| ||
Db       2941 GGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTATGGTGACCCGACTAATGGAGGCAA 3000

Qy       3093 ATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTG 3152
             |||||||||||||||||||||||||||||||| || |||||||||||||||| ||||||||
Db       3001 ATGTCAGCCATGCAAGTGCAATGGGCACGCATCACTGTGCAACACCAACACCGGCAAGTG 3060

Qy       3153 CTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCG 3212
             |||||| |||||||| || |||||||||||||||||||||||||||||||||||||||||
Db       3061 CTTCTGTACCACCAAAGGTGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCG 3120

Qy       3213 ATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTT 3272
             |||||| |||||||||||| ||||||||| || ||||| ||||| ||||||||||||||
Db       3121 ATACCAGGGAAACCCTCTCAAAGGAACATGCTACTATACCCTTCTCATTGACTATCAGTT 3180

Qy       3273 CACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTAC 3332
             |||||||| || |||||||||| || ||||| |||||||| ||||| ||||||||||||
Db       3181 CACCTTTAGCCTGTCCCAGGAAGACGACCGCTACTACACAGCCATCAACTTTGTGGCTAC 3240

Qy       3333 TCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCT 3392
             |||||| |||||||||||||||||||||||||||||||||||||||||| || ||||||||
Db       3241 TCCTGATGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAAAACTTCAACCT 3300

Qy       3393 CAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGT 3452
             |||||||||||||| |||| ||| |||| ||||||||| |||||||||| |||||||||
Db       3301 CAACATCACCTGGGCCACCAGCTTCCCAGCCGGAACCCAGACTGGAGAAGAGGTGCCTGT 3360

Qy       3453 TGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTT 3512
             |||||||||||||||||| |||| |||||||||| |||||||||||||||| ||||||||
Db       3361 TGTTTCAAAAACCAACATCAAGGAATACAAAGATAGCTTCTCTAATGAGAAATTTGATTT 3420

Qy       3513 TCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAA 3572
             ||||||||| |||| |||||||||||||||||||||||||||||||| ||||||||||||
Db       3421 TCGCAACCATCCAAACATCACTTTCTTTGTTTATGTCAGTAATTTCACTTGGCCCATCAA 3480

Qy       3573 AATTCAGGT 3581
```

```
             ||||||| |
Db       3481 AATTCAGAT 3489
```

RESULT 4
BX440935
LOCUS BX440935 1201 bp mRNA linear EST 15-MAY-2003
DEFINITION BX440935 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone CS0DF012YL19 5-PRIME, mRNA sequence.
ACCESSION BX440935
VERSION BX440935.1 GI:30781889
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 10212.f For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DF012CF10QP1&cluster=10212.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DF012CF10QP1.
FEATURES Location/Qualifiers
source 1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF012YL19"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6. vector. Library was not normalized."
ORIGIN

Query Match 25.7%; Score 925.6; DB 13; Length 1201;
Best Local Similarity 98.4%; Pred. No. 4.7e-223;
Matches 962; Conservative 4; Mismatches 8; Indels 4; Gaps 3;

```
Qy       2115 GTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATG 2174
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         57 GTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATG 116

Qy       2175 TTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCA 2234
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        117 TTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCA 176

Qy       2235 CAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCT 2294
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        177 CAATGCCCTTTTGGCTTCTCTTACAACCCTGATGAAGGTAGAATTTGTCCTTAAGCAGCT 236

Qy       2295 GCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCT 2354
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        237 GCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCT 296

Qy       2355 TCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTT 2414
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 TCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTT 356

Qy       2415 ACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACC 2474
```

```
Db        357 ACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACC 416

Qy       2475 CAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGA 2534

Db        417 CAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGA 476

Qy       2535 AAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATG 2594

Db        477 AAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATG 536

Qy       2595 TGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGT 2654

Db        537 TGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGT 596

Qy       2655 GGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGAT 2714

Db        597 GGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGAT 656

Qy       2715 GAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCA 2774

Db        657 GAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCA 716

Qy       2775 ACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGG 2834

Db        717 ACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGG 776

Qy       2835 TTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACA 2894

Db        777 TTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACA 836

Qy       2895 GCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCA 2954

Db        837 GCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCA 896

Qy       2955 CTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAA 3014

Db        897 CTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAA 956

Qy       3015 GTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGA 3074

Db        957 GTGTGAGAACYGAAC--AMAGGMAGCACTGCGAGAACTGCATATCTGGC-TCTACGGKGA 1013

Qy       3075 TCCCACCAATGGAGGGAA 3092

Db       1014 T-CCACCAATGGAGGGAA 1030
```

```
RESULT 5
BU506373
LOCUS       BU506373                964 bp    mRNA    linear   EST 12-SEP-2002
DEFINITION  AGENCOURT_10015485 NIH_MGC_94 Mus musculus cDNA clone IMAGE:6494441
            5', mRNA sequence.
ACCESSION   BU506373
VERSION     BU506373.1  GI:22812606
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 964)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: The Cepko Laboratory
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM14050  row: k  column: 18
            High quality sequence stop: 592.
FEATURES             Location/Qualifiers
     source          1. .964
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:6494441"
                     /tissue_type="retina"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_94"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 3.3 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
ORIGIN
```

Query Match 21.5%; Score 775; DB 13; Length 964;
Best Local Similarity 88.9%; Pred. No. 6.8e-185;
Matches 857; Conservative 0; Mismatches 105; Indels 2; Gaps 2;

```
Qy       1774 TGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAA 1833

Db          1 TGTCTGTGGGACACACAGTCGTCTCGATGTACCTCCTGGGAGTTGGCAACTGAAGAACAA 60

Qy       1834 GAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGAC 1893

Db         61 GCAGAAAAGTTAAAATCAGAGTGTTTTTCTAAAAGAACCCTTGACCATGACAGATGTGAC 120

Qy       1894 CAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGCAATGAC 1953

Db        121 CAGCACACAGATTGTTACAGCTGCACAGCCAATACCAATGACTGCCACTGGTGCAATGAT 180

Qy       1954 CATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAG 2013

Db        181 CACTGTGTCCCTGTGAACCACAGCTGCACAGAAGGCCAGATCTCCATTGCCAAGTATGAG 240

Qy       2014 AATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGT 2073

Db        241 AGTTGCCCCAAGGATAACCCCATGTACTACTGCAATAAGAAAACCAGCTGCAGGAGCTGT 300

Qy       2074 GCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAA 2133

Db        301 GCCCTAGACCAGAACTGCCAGTGGGAGCCCCGGAATCAAGAGTGCATCGCCCTGCCGGAA 360

Qy       2134 AATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCC 2193

Db        361 AATATCTGTGGCAATGGCTGGCATTTGGTTGGAAACTCGTGTCTGAAAATCACTACTGCT 420

Qy       2194 AAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCT 2253

Db        421 AAGGAGAATTATGACAATGCTAAATTGTCCTGTAGGAACCACAATGCCTTTTTGGCTTCC 480

Qy       2254 CTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCT 2313

Db        481 CTCACATCCCAGAAGAAGGTGGAGTTTGTCCTTAAGCAGCTTCGATTAATGCAATCATCT 540

Qy       2314 CAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCC 2373

Db        541 CAAAGTATGTCCAAGCTCACTCTGACTCCATGGGTTGGTCTTCGGAAGATCAATGTGTCT 600

Qy       2374 TACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCT 2433

Db        601 TACTGGTGCTGGGAGGATATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCT 660

Qy       2434 GAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTC-GGGGACTGAA 2492

Db        661 GAGCCCAGTGATGCTGGCTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGGATTAAA 720

Qy       2493 GGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAG 2552

Db        721 GGCTGCAACCTGCATCAACCCTCTCAATGGCAGCGTCTGTGAAAAGGCTGCAAACCACAG 780
```

```
Qy       2553 TGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAG-CATGTGGAGATTGCACCAGCG 2611
              ||| |||||||||||||||||||||||| || |||||  | |||||  | |||||  ||| 
Db        781 TGCCAAGCAGTGCCGGACACCATGTGCCCTGCGGACANCCTTGTGGCCAGTGCACTAGCA 840

Qy       2612 GCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATG 2671
              |||||| |||||||||||||||| |||||||||| |||| ||||||||||||||| | 
Db        841 CCAGCTCGGAGTGCATGTGGTGCAGTAACATGAAGCAATGTGGGGACTCCAATGCCTACG 900

Qy       2672 TGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTG 2731
              |||||||||||| ||||| |||  | || ||| |||||||||||||| ||| ||  ||| 
Db        901 TGGCCTCCTTCCNTTTTGNCCAANGNATTGAAAGGTATACGATGAGCAGCTGTCCAACTG 960

Qy       2732 AAAA 2735
              |||| 
Db        961 AAAA 964
```

RESULT 6
BG678679
LOCUS BG678679 933 bp mRNA linear EST 01-MAY-2001
DEFINITION 602624477F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4749427 5', mRNA sequence.
ACCESSION BG678679
VERSION BG678679.1 GI:13910076
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 933)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10602 row: j column: 20
High quality sequence stop: 762.
FEATURES Location/Qualifiers
source 1. .933
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4749427"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Skn4"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5kb. Library constructed by Life Technologies. Note: this is a NCI_CGAP Library."
ORIGIN

Query Match 21.2%; Score 763.2; DB 12; Length 933;
Best Local Similarity 97.3%; Pred. No. 6.6e-182;
Matches 798; Conservative 0; Mismatches 18; Indels 4; Gaps 2;

```
Qy       2775 ACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGG 2834
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db          2 ACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGG 61

Qy       2835 TTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACA 2894
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db         62 TTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACA 121

Qy       2895 GCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCA 2954
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db        122 GCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCA 181

Qy       2955 CTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAA 3014
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db        182 CTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAA 241

Qy       3015 GTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGA 3074
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db        242 GTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGA 301

Qy       3075 TCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAA 3134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db        302 TCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAA 361

Qy       3135 CACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCT 3194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db        362 CACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCT 421

Qy       3195 ATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCT 3254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db        422 ATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCT 481

Qy       3255 TCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGC 3314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db        482 TCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGC 541

Qy       3315 TATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGC 3374
              ||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||| 
Db        542 TATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTCGGACATGTTCATCAATGC 601

Qy       3375 CTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGC 3434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db        602 CTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGC 661

Qy       3435 TGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTC 3494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db        662 TGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTC 721

Qy       3495 TAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAA 3554
              |||||||||||  |||| |||||| || |||||||||||||| ||| ||||||||||| 
Db        722 TAATGAGAAGTCCGATTCTCGCAAGCA-CCAAATATCACTTTCTGTGTGTATGTCAGTAA 780

Qy       3555 TTTCACCT---GGCCCATCAAAATTCAGGTGCAAACTGAA 3591
              ||   ||  | ||   |||||||||||||| || | 
Db        781 TTGTCACTTGGGCCCATCGAAAATTCAGGTGCAAAATGGA 820
```

RESULT 7
CB520374
LOCUS CB520374 840 bp mRNA linear EST 09-JUL-2003
DEFINITION UI-M-GI0-cei-l-18-0-UI.r1 NIH_BMAP_GI0 Mus musculus cDNA clone IMAGE:6840067 5', mRNA sequence.
ACCESSION CB520374
VERSION CB520374.1 GI:29353729
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 840)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)
Seq primer: pYX-5.
FEATURES Location/Qualifiers
source 1. .840
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6840067"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_GI0"
/note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel.First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site .Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector . The library tag sequence located between the Not I site and the polyA tail is AGCGAGACAG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."
ORIGIN

Query Match 19.4%; Score 699.6; DB 14; Length 840;
Best Local Similarity 90.1%; Pred. No. 8.6e-166;
Matches 758; Conservative 0; Mismatches 82; Indels 1; Gaps 1;

```
Qy       1929 CAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGG 1988
             |||||||||||||||||||||||| || ||||||||   |||||||||||| |||||||
Db          1 CAATGACTGCCACTGGTGCAATGATCACTGTGTCCCTGTGAACCACAGCTGCACAGAAGG 60

Qy       1989 CCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAA 2048
             |||||||||||||   | |||||||| |||||||||||||||||| |||||||||| ||
Db         61 CCAGATCTCCATTGCCAAGTATGAGAGTTGCCCCAAGGATAACCCCATGTACTACTGCAA 120

Qy       2049 CAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAA 2108
             ||||| ||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        121 TAAGAAAACCAGCTGCAGGAGCTGTGCCCTAGACCAGAACTGCCAGTGGGAGCCCCGGAA 180

Qy       2109 TCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAA 2168
             ||| |||||||| |||||||| |||||||||||||||| ||||||||||||||||||||||
Db        181 TCAAGAGTGCATCGCCCTGCCGGAAAATATCTGTGGCAATGGCTGGCATTTGGTTGGAAA 240

Qy       2169 CTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAG 2228
             ||| ||| |||||| |||||||| ||||||||||||||||||||||||||||||||||||
Db        241 CTCGTGTCTGAAAATCACTACTGCTAAGGAGAATTATGACAATGCTAAATTGTCCTGTAG 300

Qy       2229 GAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAA 2288
             |||||||||||| |||||||||| || ||| |||||||||||||| || ||||||||||
Db        301 GAACCACAATGCCTTTTTGGCTTCCCTCACATCCCAGAAGAAGGTGGAGTTTGTCCTTAA 360

Qy       2289 GCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGT 2348
             |||||| ||| ||||||| |||||||| || |||||||||||||  | || ||||||||
Db        361 GCAGCTTCGATTAATGCAATCATCTCAAAGTATGTCCAAGCTCACTCTGACTCCATGGGT 420

Qy       2349 CGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAA 2408
             || |||||||||||||||||| |||||||||||||| |||||||| ||||| |||||||
Db        421 TGGTCTTCGGAAGATCAATGTGTCTTACTGGTGCTGGGAGGATATGTCTCCATTCACAAA 480

Qy       2409 TAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATC 2468
             |||||| || |||||||||| |||||||||||||||||||| |||||||| || || ||
Db        481 TAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGTGATGCTGGCTTCTGTGGGATCTTGTC 540

Qy       2469 AGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGT 2528
             ||| || |||||||||| | |||||||||||||||||||| |||||||| || ||
Db        541 AGAGCCTAGTACTCGGGGATTAAAGGCTGCAACCTGCATCAACCCTCTCAATGGCAGCGT 600

Qy       2529 CTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGAC 2588
             |||||||||||||||| |||||||||| |||||||||||||||||||||||||| || ||||
Db        601 CTGTGAAAGGCCTGCANACCACAGTGCCAAGCAGTGCCGGACACCATGTGCCCTGCGGAC 660

Qy       2589 AGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCA 2648
             ||| ||||| || ||||| ||| ||||||| |||||||||||||||     ||||||||||
Db        661 AGCGTGTGGCGAGTGCACTAGCAGCAGCTCGGAGTGCATGTGGTGCNAGTACATGAAGCA 720

Qy       2649 GTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTA 2708
             |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db        721 GTGTGTGGACTCCAATGCCTACGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTA 780

Qy       2709 TACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTT 2768
             |||||||||| |||||| |||| | |||| || |||||||||||||| || ||||||||
Db        781 TACGATGAGCAGCTGCCCACCTG-ANATTGCTCTGGCTACTGTACCTGCAGCCATTGCTT 839

Qy       2769 G 2769
             |
Db        840 G 840
```

RESULT 8
BM783739
LOCUS BM783739 681 bp mRNA linear EST 05-MAR-2002
DEFINITION K-EST0061730 S5SNU484 Homo sapiens cDNA clone S5SNU484-34-H05 5', mRNA sequence.
ACCESSION BM783739
VERSION BM783739.1 GI:19131971
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 681)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 34 row: H column: 05
High quality sequence stop: 681.
FEATURES Location/Qualifiers
source 1. .681
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S5SNU484-34-H05"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/cell_line="SNU-484"
/lab_host="Top10F'"
/clone_lib="S5SNU484"
/note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI; Site_2: NotI; The poly (A)+ RNA was decapped with tabacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation

method."
ORIGIN

Query Match 18.9%; Score 681; DB 12; Length 681;
Best Local Similarity 100.0%; Pred. No. 4e-161;
Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2290 CAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTC 2349
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTC 60

Qy       2350 GGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAAT 2409
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAAT 120

Qy       2410 AGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCA 2469
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCA 180

Qy       2470 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTC 2529
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTC 240

Qy       2530 TGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACA 2589
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACA 300

Qy       2590 GCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 2649
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 360

Qy       2650 TGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 2709
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 420

Qy       2710 ACGATGAGCACCTGCCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTG 2769
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACGATGAGCACCTGCCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTG 480

Qy       2770 GAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATA 2829
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATA 540

Qy       2830 GAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTAT 2889
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTAT 600

Qy       2890 CCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 2949
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 660

Qy       2950 ATTCACTGTCCAGCTTGCCAA 2970
             |||||||||||||||||||||
Db        661 ATTCACTGTCCAGCTTGCCAA 681
```

RESULT 9
BM783788
LOCUS BM783788 683 bp mRNA linear EST 05-MAR-2002
DEFINITION K-EST0061794 S5SNU484 Homo sapiens cDNA clone S5SNU484-36-E09 5', mRNA sequence.
ACCESSION BM783788
VERSION BM783788.1 GI:19132020
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 683)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 36 row: E column: 09
High quality sequence stop: 683.
FEATURES Location/Qualifiers
source 1. .683
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S5SNU484-36-E09"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/cell_line="SNU-484"
/lab_host="Top10F'"
/clone_lib="S5SNU484"
/note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI; Site_2: NotI; The poly (A)+ RNA was decapped with tabacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method."
ORIGIN

Query Match 18.9%; Score 680.4; DB 12; Length 683;
Best Local Similarity 99.7%; Pred. No. 5.7e-161;
Matches 681; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       2290 CAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTC 2349
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTC 60

Qy       2350 GGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAAT 2409
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAAT 120

Qy       2410 AGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCA 2469
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCA 180

Qy       2470 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTC 2529
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTC 240

Qy       2530 TGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACA 2589
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACA 300

Qy       2590 GCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 2649
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 360

Qy       2650 TGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 2709
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 420

Qy       2710 ACGATGAGCACCTGCCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTG 2769
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACGATGAGCACCTGCCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTG 480
```

```
Qy       2770 GAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATA 2829
             ||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db        481 GAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGNCATACTGGCAAAGGGAAATGCATA 540

Qy       2830 GAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTAT 2889
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTAT 600

Qy       2890 CCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 2949
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 660

Qy       2950 ATTCACTGTCCAGCTTGCCAATG 2972
             |||||||||||||||||||||||
Db        661 ATTCACTGTCCAGCTTGCCAATG 683


RESULT 10
BM783258
LOCUS       BM783258                680 bp    mRNA    linear   EST 05-MAR-2002
DEFINITION  K-EST0061157 S5SNU484 Homo sapiens cDNA clone S5SNU484-31-H08 5',
            mRNA sequence.
ACCESSION   BM783258
VERSION     BM783258.1  GI:19131490
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 680)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 31  row: H  column: 08
            High quality sequence stop: 680.
FEATURES             Location/Qualifiers
     source          1. .680
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="S5SNU484-31-H08"
                     /sex="M"
                     /tissue_type="Stomach"
                     /cell_type="Epithelial"
                     /cell_line="SNU-484"
                     /lab_host="Top10F'"
                     /clone_lib="S5SNU484"
                     /note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI;
                     Site_2: NotI; The poly (A)+ RNA was decapped with tabacco
                     acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
                     including EcoRI site by treatment of T4 RNA ligase. The
                     first strand cDNA was synthesized from oligo dT-selected
                     mRNA by priming with dT-tailed vector. The dT-tailed
                     vector was adjusted to have about 60nt. The cDNA vector
                     was circularized with E. coli DNA ligase after digestion
                     of EcoRI which site is also included in vector. An RNA
                     strand converted to a DNA strand by Okayama-Berg method.
                     The obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation
                     method."
ORIGIN

  Query Match           18.9%;  Score 680;  DB 12;  Length 680;
  Best Local Similarity 100.0%;  Pred. No. 7.2e-161;
  Matches  680;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2290 CAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTC 2349
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTC 60

Qy       2350 GGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAAT 2409
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAAT 120

Qy       2410 AGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCA 2469
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 AGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCA 180

Qy       2470 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTC 2529
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTC 240

Qy       2530 TGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACA 2589
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 TGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACA 300

Qy       2590 GCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 2649
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 GCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 360

Qy       2650 TGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 2709
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 TGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 420

Qy       2710 ACGATGAGCACCTGCCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTG 2769
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 ACGATGAGCACCTGCCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTG 480

Qy       2770 GAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATA 2829
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 GAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATA 540

Qy       2830 GAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTAT 2889
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 GAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTAT 600

Qy       2890 CCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 2949
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 CCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 660

Qy       2950 ATTCACTGTCCAGCTTGCCA 2969
             ||||||||||||||||||||
Db         661 ATTCACTGTCCAGCTTGCCA 680


RESULT 11
BM783773
LOCUS       BM783773                680 bp    mRNA    linear   EST 05-MAR-2002
DEFINITION  K-EST0061775 S5SNU484 Homo sapiens cDNA clone S5SNU484-36-D02 5',
            mRNA sequence.
ACCESSION   BM783773
VERSION     BM783773.1  GI:19132005
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 680)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
```

Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 36 row: D column: 02
High quality sequence stop: 680.

```
FEATURES             Location/Qualifiers
     source          1. .680
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="S5SNU484-36-D02"
                     /sex="M"
                     /tissue_type="Stomach"
                     /cell_type="Epithelial"
                     /cell_line="SNU-484"
                     /lab_host="Top10F'"
                     /clone_lib="S5SNU484"
                     /note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI;
                     Site_2: NotI; The poly (A)+ RNA was decapped with tabacco
                     acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
                     including EcoRI site by treatment of T4 RNA ligase. The
                     first strand cDNA was synthesized from oligo dT-selected
                     mRNA by priming with dT-tailed vector. The dT-tailed
                     vector was adjusted to have about 60nt. The cDNA vector
                     was circularized with E. coli DNA ligase after digestion
                     of EcoRI which site is also included in vector. An RNA
                     strand converted to a DNA strand by Okayama-Berg method.
                     The obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation
                     method."
ORIGIN

  Query Match           18.9%;  Score 680;  DB 12;  Length 680;
  Best Local Similarity  100.0%;  Pred. No. 7.2e-161;
  Matches  680;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2290 CAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTC 2349
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTC 60

Qy       2350 GGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAAT 2409
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAAT 120

Qy       2410 AGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCA 2469
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCA 180

Qy       2470 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTC 2529
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTC 240

Qy       2530 TGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACA 2589
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACA 300

Qy       2590 GCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 2649
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 360

Qy       2650 TGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 2709
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 420

Qy       2710 ACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTG 2769
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTG 480

Qy       2770 GAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATA 2829
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATA 540

Qy       2830 GAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTAT 2889
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTAT 600

Qy       2890 CCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 2949
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 660

Qy       2950 ATTCACTGTCCAGCTTGCCA 2969
             ||||||||||||||||||||
Db        661 ATTCACTGTCCAGCTTGCCA 680


RESULT 12
BM783718
LOCUS       BM783718                677 bp    mRNA    linear   EST 05-MAR-2002
DEFINITION  K-EST0061702 S5SNU484 Homo sapiens cDNA clone S5SNU484-34-F01 5',
            mRNA sequence.
ACCESSION   BM783718
VERSION     BM783718.1  GI:19131950
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 677)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 34  row: F  column: 01
            High quality sequence stop: 677.
FEATURES             Location/Qualifiers
     source          1. .677
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="S5SNU484-34-F01"
                     /sex="M"
                     /tissue_type="Stomach"
                     /cell_type="Epithelial"
                     /cell_line="SNU-484"
                     /lab_host="Top10F'"
                     /clone_lib="S5SNU484"
                     /note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI;
                     Site_2: NotI; The poly (A)+ RNA was decapped with tabacco
                     acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
                     including EcoRI site by treatment of T4 RNA ligase. The
                     first strand cDNA was synthesized from oligo dT-selected
                     mRNA by priming with dT-tailed vector. The dT-tailed
                     vector was adjusted to have about 60nt. The cDNA vector
                     was circularized with E. coli DNA ligase after digestion
                     of EcoRI which site is also included in vector. An RNA
                     strand converted to a DNA strand by Okayama-Berg method.
                     The obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation
                     method."
ORIGIN

  Query Match           18.8%;  Score 677;  DB 12;  Length 677;
  Best Local Similarity  100.0%;  Pred. No. 4.2e-160;
  Matches  677;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2290 CAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTC 2349
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTC 60
```

```
Qy       2350 GGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAAT 2409
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAAT 120

Qy       2410 AGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCA 2469
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCA 180

Qy       2470 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTC 2529
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTC 240

Qy       2530 TGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACA 2589
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACA 300

Qy       2590 GCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 2649
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 360

Qy       2650 TGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 2709
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 420

Qy       2710 ACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTG 2769
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTG 480

Qy       2770 GAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATA 2829
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATA 540

Qy       2830 GAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTAT 2889
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTAT 600

Qy       2890 CCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 2949
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 660

Qy       2950 ATTCACTGTCCAGCTTG 2966
             ||||||||||||||||
Db        661 ATTCACTGTCCAGCTTG 677
```

RESULT 13
BM783821

```
LOCUS       BM783821                677 bp    mRNA    linear   EST 05-MAR-2002
DEFINITION  K-EST0061830 S5SNU484 Homo sapiens cDNA clone S5SNU484-36-H09 5',
            mRNA sequence.
ACCESSION   BM783821
VERSION     BM783821.1  GI:19132053
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 677)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 36  row: H  column: 09
            High quality sequence stop: 677.
```

```
FEATURES             Location/Qualifiers
     source          1. .677
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="S5SNU484-36-H09"
                     /sex="M"
                     /tissue_type="Stomach"
                     /cell_type="Epithelial"
                     /cell_line="SNU-484"
                     /lab_host="Top10F'"
                     /clone_lib="S5SNU484"
                     /note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI;
                     Site_2: NotI; The poly (A)+ RNA was decapped with tabacco
                     acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
                     including EcoRI site by treatment of T4 RNA ligase. The
                     first strand cDNA was synthesized from oligo dT-selected
                     mRNA by priming with dT-tailed vector. The dT-tailed
                     vector was adjusted to have about 60nt. The cDNA vector
                     was circularized with E. coli DNA ligase after digestion
                     of EcoRI which site is also included in vector. An RNA
                     strand converted to a DNA strand by Okayama-Berg method.
                     The obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation
                     method."
ORIGIN
```

```
  Query Match             18.7%;  Score 674.4;  DB 12;  Length 677;
  Best Local Similarity   99.7%;  Pred. No. 1.9e-159;
  Matches  675;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2290 CAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTC 2349
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTC 60

Qy       2350 GGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAAT 2409
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAAT 120

Qy       2410 AGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCA 2469
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCA 180

Qy       2470 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTC 2529
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTC 240

Qy       2530 TGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACA 2589
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACA 300

Qy       2590 GCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 2649
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 360

Qy       2650 TGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 2709
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 420

Qy       2710 ACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTG 2769
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTG 480

Qy       2770 GAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATA 2829
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATA 540

Qy       2830 GAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTAT 2889
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGNCAGCCCCTACAGGAAATTTCTAT 600

Qy       2890 CCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 2949
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        601 CCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 660

Qy       2950 ATTCACTGTCCAGCTTG 2966
              ||||||||||||||||
Db        661 ATTCACTGTCCAGCTTG 677


RESULT 14
BM783685
LOCUS       BM783685                673 bp    mRNA    linear   EST 05-MAR-2002
DEFINITION  K-EST0061659 S5SNU484 Homo sapiens cDNA clone S5SNU484-34-B06 5',
            mRNA sequence.
ACCESSION   BM783685
VERSION     BM783685.1  GI:19131917
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 673)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 34  row: B  column: 06
            High quality sequence stop: 673.
FEATURES             Location/Qualifiers
     source          1. .673
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="S5SNU484-34-B06"
                     /sex="M"
                     /tissue_type="Stomach"
                     /cell_type="Epithelial"
                     /cell_line="SNU-484"
                     /lab_host="Top10F'"
                     /clone_lib="S5SNU484"
                     /note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI;
                     Site_2: NotI; The poly (A)+ RNA was decapped with tabacco
                     acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
                     including EcoRI site by treatment of T4 RNA ligase. The
                     first strand cDNA was synthesized from oligo dT-selected
                     mRNA by priming with dT-tailed vector. The dT-tailed
                     vector was adjusted to have about 60nt. The cDNA vector
                     was circularized with E. coli DNA ligase after digestion
                     of EcoRI which site is also included in vector. An RNA
                     strand converted to a DNA strand by Okayama-Berg method.
                     The obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation
                     method."
ORIGIN

  Query Match             18.7%;  Score 673;  DB 12;  Length 673;
  Best Local Similarity   100.0%;  Pred. No. 4.3e-159;
  Matches  673;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2290 CAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTC 2349
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTC 60

Qy       2350 GGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAAT 2409
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAAT 120

Qy       2410 AGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCA 2469
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCA 180

Qy       2470 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTC 2529
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTC 240

Qy       2530 TGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACA 2589
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACA 300

Qy       2590 GCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 2649
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 360

Qy       2650 TGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 2709
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 420

Qy       2710 ACGATGAGCACCTGCCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTG 2769
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACGATGAGCACCTGCCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTG 480

Qy       2770 GAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATA 2829
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATA 540

Qy       2830 GAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTAT 2889
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTAT 600

Qy       2890 CCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 2949
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 660

Qy       2950 ATTCACTGTCCAG 2962
              |||||||||||||
Db        661 ATTCACTGTCCAG 673


RESULT 15
BX431593
LOCUS       BX431593                962 bp    mRNA    linear   EST 15-MAY-2003
DEFINITION  BX431593 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
            CS0DF012YL19 5-PRIME, mRNA sequence.
ACCESSION   BX431593
VERSION     BX431593.1  GI:30781048
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 962)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 10212.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0BAG017ZB02_CS01546_1&cluster=10212.f.
            Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0BAG017ZB02_CS01546_1.
FEATURES             Location/Qualifiers
     source          1. .962
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
```

/db_xref="taxon:9606"
/clone="CS0DF012YL19"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 18.6%; Score 669.8; DB 13; Length 962;
Best Local Similarity 98.4%; Pred. No. 3.4e-158;
Matches 697; Conservative 0; Mismatches 8; Indels 3; Gaps 2;

```
Qy       2891 CACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATA-CAACTGGTCTTTC 2949
             ||||||||||||| | | |  ||||||||||||||||||||||| ||||||||||||
Db          7 CACAGCCCCTGCT--AATTCCAGATGTGTCTAGAGGACAGCAGATACCAACTGGTCTTTC 64

Qy       2950 ATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGT 3009
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 ATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGT 124

Qy       3010 GAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTAC 3069
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 GAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTAC 184

Qy       3070 GGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTG 3129
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 GGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTG 244

Qy       3130 TGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGC 3189
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 TGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGC 304

Qy       3190 CAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTAT 3249
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 CAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTAT 364

Qy       3250 ACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTAC 3309
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 ACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTAC 424

Qy       3310 ACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATC 3369
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        425 ACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATC 484

Qy       3370 AATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACC 3429
             |||||||||||||||||||||||||||||||  |||||||||||||||||||||||||||
Db        485 AATGCCTCCAAGAATTTCAACCTCAACATCACCNTGGCTGCCAGTTTCTCAGCTGGAACC 544

Qy       3430 CAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGT 3489
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        545 CAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGT 604

Qy       3490 TTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTC 3549
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        605 TTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTC 664

Qy       3550 AGTAATTTCACCTGGCCCATCAAAATTCAGGTGCAAACTGAACAATGA 3597
             ||||| |||||||||||||||||||||||||||||||||||||||||
Db        665 AGTAAATTCACCTGGCCCATCAAAATTCAGGTGCAAACTGAACAATGA 712
```

Search completed: March 5, 2004, 12:07:51
Job time : 5707.24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2004, 16:06:34 ; Search time 847.6 Seconds
(without alignments)
18028.269 Million cell updates/sec

Title: US-09-787-097-1
Perfect score: 3597
Sequence: 1 atggtggccgcagcggcggc.....aggtgcaaactgaacaatga 3597

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	3597	100.0	3597	3	AAZ52274	Aaz52274 Human sol
	2	3597	100.0	3597	5	AAS72657	Aas72657 DNA encod
	3	3597	100.0	3597	9	ADD70997	Add70997 Human att
	4	3556.4	98.9	4068	3	AAZ52275	Aaz52275 Human mem
	5	3494.4	97.1	4072	3	AAZ91921	Aaz91921 Human mah
	6	3487	96.9	3819	3	AAZ52277	Aaz52277 Human sol
	7	3476.8	96.7	8590	3	AAZ91920	Aaz91920 Human mah
	8	3469.4	96.5	4290	3	AAZ52276	Aaz52276 Human mem
	9	3469.4	96.5	8561	5	AAS72659	Aas72659 DNA encod
	10	2909	80.9	8739	9	ADB53249	Adb53249 Primary r
	11	2843.6	79.1	8827	3	AAZ91916	Aaz91916 Wild type
	12	2790.4	77.6	3490	5	AAS72660	Aas72660 DNA encod
	13	2028.8	56.4	2625	3	AAZ91922	Aaz91922 Human mah
	14	1428.8	39.7	2419	3	AAZ91917	Aaz91917 Murine ma
	15	1341.6	37.3	6373	3	AAZ91919	Aaz91919 Human mah
	16	1212.2	33.7	4134	7	ABZ58710	Abz58710 Murine ho
	17	1203.2	33.5	4140	7	ABZ58709	Abz58709 Human hom
	18	1129.4	31.4	3609	6	ABS63434	Abs63434 Human cDN
	19	873.2	24.3	3165	5	AAS78781	Aas78781 DNA encod
	20	873.2	24.3	3313	5	AAS76294	Aas76294 DNA encod
	21	817.8	22.7	1012	2	AAX84568	Aax84568 Biorhythm
	22	574	16.0	6517	4	AAH98748	Aah98748 Human EST
	23	536	14.9	6532	5	AAS86766	Aas86766 DNA encod
	24	463	12.9	1051	3	AAZ91918	Aaz91918 Murine ma
	25	421.8	11.7	671	5	AAS72658	Aas72658 DNA encod
	26	418.4	11.6	1654	7	ABX34730	Abx34730 Human mdd
	27	413.6	11.5	1383	6	AAL54064	Aal54064 cDNA of h
	28	332	9.2	668	4	AAS02409	Aas02409 Human sec
	29	278	7.7	292	3	AAA45505	Aaa45505 Human sec
c	30	277.4	7.7	481	4	AAL14686	Aal14686 Human bre
c	31	266.4	7.4	482	4	AAL23548	Aal23548 Human bre
	32	253.8	7.1	982	6	ABL39784	Abl39784 Human NS
	33	250.8	7.0	531	4	AAI43425	Aai43425 Probe #12
	34	250.8	7.0	531	6	ABS11523	Abs11523 Human gen
c	35	246	6.8	246	6	ABK38928	Abk38928 cDNA enco
c	36	246	6.8	246	7	ACA11257	Aca11257 Human lun
c	37	246	6.8	246	7	ACA02443	Aca02443 Lung canc
	38	236	6.6	657	7	ABZ58714	Abz58714 Upregulat
	39	234.4	6.5	207433	5	ABZ72040	Abz72040 Gene 216
	40	234.4	6.5	207433	7	ABX74891	Abx74891 BAC1098L2
	41	202.8	5.6	90050	3	AAZ91925	Aaz91925 Wild type
	42	193	5.4	193	4	AAI56405	Aai56405 Probe #25
	43	193	5.4	193	6	ABS23881	Abs23881 Human gen
c	44	182	5.1	182	4	AAI24099	Aai24099 Probe #14
c	45	182	5.1	182	4	ABA69223	Aba69223 Human foe

ALIGNMENTS

RESULT 1
AAZ52274
ID AAZ52274 standard; cDNA; 3597 BP.
XX
AC AAZ52274;
XX
DT 18-JUL-2000 (first entry)
XX
DE Human soluble attractin-1 cDNA.
XX
KW Human; soluble attractin-1; immune response; macrophage; monocyte;
KW T cell; immunostimulant; immunosuppressed patient; cancer;
KW immunodeficiency syndrome; transplant; autoimmune disease; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1. .3597
FT /*tag= a
FT /product= "Soluble attractin-1"
XX
PN WO200015651-A1.
XX
PD 23-MAR-2000.
XX
PF 14-SEP-1999; 99WO-US020948.
XX
PR 14-SEP-1998; 98US-0100137P.
XX
PA (DAND) DANA FARBER CANCER INST INC.
XX
PI Duke-Cohan JS, Schlossman SF;
XX
DR WPI; 2000-271373/23.
DR P-PSDB; AAY70689.
XX
PT Isolated nucleic acids encoding human attractin polypeptides useful for
PT enhancing immune responses.
XX
PS Claim 3; Fig 8; 120pp; English.
XX
CC The patent discloses four forms of human attractin polypeptides which
CC enhance immune response by promoting macrophage and monocyte spreading in
CC the presence of T cells. These include soluble attractin-1 and -2 and
CC membrane attractin-1 and -2. These various forms of attractin are encoded
CC by alternatively spliced mRNA molecule transcribed from a single gene.

CC The present sequence is a cDNA encoding soluble attractin-1 obtained from
CC activated T cell and human foetal liver libraries. This sequence can be
CC used to enhance immune response in immunosuppressed patients such as
CC those undergoing chemo- and radio-therapy treatment for cancer or those
CC suffering from common variable immunodeficiency syndrome. The proteins
CC may also be used to screen modulators (agonists and antagonists) of
CC immune responses which may also be used to regulate immune reactions.
CC Attractin antibodies can be used to inhibit immune response in transplant
CC recipients or patients afflicted with autoimmune disease
XX
SQ Sequence 3597 BP; 951 A; 822 C; 898 G; 926 T; 0 U; 0 Other;

Query Match 100.0%; Score 3597; DB 3; Length 3597;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 ATGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 60

Qy   61 GCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGC 120

Qy  121 ACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 ACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGC 180

Qy  181 CGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATAC 240

Qy  241 AAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 AAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTC 300

Qy  301 AATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 AATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATT 360

Qy  361 TATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 TATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAG 420

Qy  421 ACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 ACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCT 480

Qy  481 GCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 GCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGC 540

Qy  541 TCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 TCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCT 600

Qy  601 GAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 GAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCT 660

Qy  661 CATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 CATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAG 720

Qy  721 GGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATAT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 GGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATAT 780

Qy  781 TCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 TCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGG 840

Qy  841 GTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 GTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTT 900

Qy  901 GCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 GCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGT 960

Qy  961 CATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 CATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACT 1020

Qy 1021 GGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 GGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTG 1080

Qy 1081 ACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 ACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTG 1140

Qy 1141 AAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATA 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 AAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATA 1200

Qy 1201 AGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 AGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGT 1260

Qy 1261 GCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTA 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 GCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTA 1320

Qy 1321 TACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTC 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1321 TACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTC 1380

Qy 1381 TACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGT 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1381 TACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGT 1440

Qy 1441 TACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACAC 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1441 TACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACAC 1500

Qy 1501 AATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGAC 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1501 AATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGAC 1560

Qy 1561 ATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATTCCACCATGATGTCAACAGAT 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1561 ATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATTCCACCATGATGTCAACAGAT 1620

Qy 1621 TTGGCCATTCCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGT 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1621 TTGGCCATTCCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGT 1680

Qy 1681 CTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAA 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1681 CTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAA 1740

Qy 1741 GCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAG 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1741 GCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAG 1800

Qy 1801 TGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTT 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1801 TGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTT 1860

Qy 1861 TCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACA 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1861 TCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACA 1920

Qy 1921 GCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGC 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1921 GCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGC 1980
Qy       1981 TCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTAC 2040
Db       1981 TCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTAC 2040
Qy       2041 TACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAG 2100
Db       2041 TACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAG 2100
Qy       2101 CCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTG 2160
Db       2101 CCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTG 2160
Qy       2161 GTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTG 2220
Db       2161 GTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTG 2220
Qy       2221 TTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTT 2280
Db       2221 TTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTT 2280
Qy       2281 GTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACC 2340
Db       2281 GTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACC 2340
Qy       2341 CCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCA 2400
Db       2341 CCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCA 2400
Qy       2401 TTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGA 2460
Db       2401 TTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGA 2460
Qy       2461 ATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAAT 2520
Db       2461 ATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAAT 2520
Qy       2521 GGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCC 2580
Db       2521 GGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCC 2580
Qy       2581 TTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAAC 2640
Db       2581 TTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAAC 2640
Qy       2641 ATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATG 2700
Db       2641 ATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATG 2700
Qy       2701 GAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGT 2760
Db       2701 GAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGT 2760
Qy       2761 CATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGG 2820
Db       2761 CATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGG 2820
Qy       2821 AAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGA 2880
Db       2821 AAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGA 2880
Qy       2881 AATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAAC 2940
Db       2881 AATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAAC 2940
Qy       2941 TGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAG 3000
Db       2941 TGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAG 3000
Qy       3001 AGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCT 3060
Db       3001 AGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCT 3060
```

```
Qy       3061 GGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCAC 3120
Db       3061 GGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCAC 3120
Qy       3121 GCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGG 3180
Db       3121 GCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGG 3180
Qy       3181 GACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACA 3240
Db       3181 GACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACA 3240
Qy       3241 TGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGAT 3300
Db       3241 TGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGAT 3300
Qy       3301 CGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGAC 3360
Db       3301 CGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGAC 3360
Qy       3361 ATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCA 3420
Db       3361 ATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCA 3420
Qy       3421 GCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTAC 3480
Db       3421 GCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTAC 3480
Qy       3481 AAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTT 3540
Db       3481 AAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTT 3540
Qy       3541 GTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGGTGCAAACTGAACAATGA 3597
Db       3541 GTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGGTGCAAACTGAACAATGA 3597
```

RESULT 2
AAS72657
ID AAS72657 standard; cDNA; 3597 BP.
XX
AC AAS72657;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #8461.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG08470.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

XX
PS Claim 1; SEQ ID NO 8461; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 3597 BP; 951 A; 822 C; 898 G; 926 T; 0 U; 0 Other;

Query Match 100.0%; Score 3597; DB 5; Length 3597;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 60

Qy         61 GCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGC 120

Qy        121 ACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGC 180

Qy        181 CGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATAC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATAC 240

Qy        241 AAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTC 300

Qy        301 AATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATT 360

Qy        361 TATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAG 420

Qy        421 ACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCT 480

Qy        481 GCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGC 540

Qy        541 TCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCT 600

Qy        601 GAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        601 GAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCT 660

Qy        661 CATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAG 720

Qy        721 GGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATAT 780

Qy        781 TCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGG 840

Qy        841 GTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTT 900

Qy        901 GCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGT 960

Qy        961 CATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACT 1020

Qy       1021 GGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTG 1080

Qy       1081 ACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTG 1140

Qy       1141 AAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATA 1200

Qy       1201 AGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGT 1260

Qy       1261 GCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTA 1320

Qy       1321 TACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTC 1380

Qy       1381 TACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGT 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGT 1440

Qy       1441 TACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACAC 1500

Qy       1501 AATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGAC 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 AATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGAC 1560

Qy       1561 ATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATTCCACCATGATGTCAACAGAT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 ATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATTCCACCATGATGTCAACAGAT 1620

Qy       1621 TTGGCCATTCCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGT 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TTGGCCATTCCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGT 1680

Qy       1681 CTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAA 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAA 1740
```

```
Qy       1741 GCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAG 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAG 1800

Qy       1801 TGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTT 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 TGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTT 1860

Qy       1861 TCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACA 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACA 1920

Qy       1921 GCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGC 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGC 1980

Qy       1981 TCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTAC 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTAC 2040

Qy       2041 TACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAG 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAG 2100

Qy       2101 CCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTG 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 CCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTG 2160

Qy       2161 GTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTG 2220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 GTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTG 2220

Qy       2221 TTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTT 2280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 TTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTT 2280

Qy       2281 GTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACC 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 GTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACC 2340

Qy       2341 CCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCA 2400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 CCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCA 2400

Qy       2401 TTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGA 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGA 2460

Qy       2461 ATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAAT 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 ATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAAT 2520

Qy       2521 GGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCC 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 GGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCC 2580

Qy       2581 TTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAAC 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 TTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAAC 2640

Qy       2641 ATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATG 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 ATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATG 2700

Qy       2701 GAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGT 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGT 2760

Qy       2761 CATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGG 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 CATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGG 2820

Qy       2821 AAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGA 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 AAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGA 2880

Qy       2881 AATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAAC 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 AATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAAC 2940

Qy       2941 TGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAG 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 TGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAG 3000

Qy       3001 AGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCT 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 AGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCT 3060

Qy       3061 GGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCAC 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 GGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCAC 3120

Qy       3121 GCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGG 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 GCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGG 3180

Qy       3181 GACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACA 3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 GACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACA 3240

Qy       3241 TGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGAT 3300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 TGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGAT 3300

Qy       3301 CGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGAC 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 CGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGAC 3360

Qy       3361 ATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCA 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 ATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCA 3420

Qy       3421 GCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTAC 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 GCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTAC 3480

Qy       3481 AAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTT 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3481 AAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTT 3540

Qy       3541 GTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGGTGCAAACTGAACAATGA 3597
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3541 GTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGGTGCAAACTGAACAATGA 3597
```

RESULT 3
ADD70997
ID ADD70997 standard; DNA; 3597 BP.
XX
AC ADD70997;
XX
DT 15-JAN-2004 (first entry)
XX
DE Human attractin gene SEQ ID NO:1.
XX
KW liver cancer; chronic hepatitis; cirrhosis; liver disease; hepatotropic;
KW cytostatic; gene therapy; human; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2003061564-A2.
XX
PD 31-JUL-2003.
XX
PF 20-DEC-2002; 2002WO-US040718.

XX
PR 21-DEC-2001; 2001US-0341815P.
PR 31-DEC-2001; 2001US-0343185P.
XX
PA (GENE-) GENE LOGIC INC.
PA (LGBI-) LG BIOMEDICAL INST.
XX
PI Koh SS, Liu Q, Chung H, Zeng W, Lee B, Yeramilli S, Song SY;
XX
DR WPI; 2003-663343/62.
XX
PT Diagnosing liver cancer cells, useful for treating liver cancer
PT associated with chronic hepatitis or cirrhosis comprises detecting the
PT level of expression in a tissue sample of one or more genes associated
PT with cancerous liver tissues.
XX
PS Claim 1; SEQ ID NO 1; 176pp; English.
XX
CC The present invention describes a method for diagnosing liver cancer
CC cells comprising detecting the level of expression in a tissue sample of
CC one or more genes given in the specification (see ADD70997 to ADD71105),
CC where differential expression of the genes is indicative of liver cancer.
CC Also described: (1) detecting the progression of liver cancer in a
CC patient; (2) monitoring the treatment of a patient with liver cancer; (3)
CC treating a patient with liver cancer; (4) typing a liver disease in a
CC patient; (5) detecting the presence or progression of liver cancer in a
CC patient with chronic hepatitis or cirrhosis; (6) differentiating liver
CC cancer related to chronic hepatitis from liver cancer related to
CC cirrhosis; (7) screening for an agent capable of modulating the onset or
CC progression of liver cancer; (8) a composition comprising at least two
CC oligonucleotides comprising a sequence that specifically hybridises to
CC any of the genes; (9) a solid support comprising the at least two
CC oligonucleotides; (10) a computer system comprising a database containing
CC information identifying the level in liver tissue of a set of genes; (11)
CC a method for using the computer system to present information identifying
CC the expression level in tissue or cell of any of the genes; and (12) a
CC therapeutic agent for slowing or halting the progression of liver cancer.
CC The methods are useful for treating liver cancer associated with chronic
CC hepatitis or cirrhosis. The present sequence represents a specifically
CC claimed human gene sequence which is used in the exemplification of the
CC present invention.
XX
SQ Sequence 3597 BP; 951 A; 822 C; 898 G; 926 T; 0 U; 0 Other;

Query Match 100.0%; Score 3597; DB 9; Length 3597;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 60

Qy         61 GCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGC 120

Qy        121 ACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGC 180

Qy        181 CGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATAC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATAC 240

Qy        241 AAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTC 300

Qy        301 AATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATT 360

Qy        361 TATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAG 420

Qy        421 ACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCT 480

Qy        481 GCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGC 540

Qy        541 TCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCT 600

Qy        601 GAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCT 660

Qy        661 CATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAG 720

Qy        721 GGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATAT 780

Qy        781 TCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGG 840

Qy        841 GTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTT 900

Qy        901 GCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGT 960

Qy        961 CATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACT 1020

Qy       1021 GGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTG 1080

Qy       1081 ACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTG 1140

Qy       1141 AAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATA 1200

Qy       1201 AGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGT 1260

Qy       1261 GCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTA 1320

Qy       1321 TACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTC 1380

Qy       1381 TACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGT 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGT 1440

Qy       1441 TACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1441 TACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACAC 1500
Qy       1501 AATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGAC 1560
Db       1501 AATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGAC 1560
Qy       1561 ATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATTCCACCATGATGTCAACAGAT 1620
Db       1561 ATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATTCCACCATGATGTCAACAGAT 1620
Qy       1621 TTGGCCATTCCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGT 1680
Db       1621 TTGGCCATTCCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGT 1680
Qy       1681 CTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAA 1740
Db       1681 CTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAA 1740
Qy       1741 GCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAG 1800
Db       1741 GCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAG 1800
Qy       1801 TGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTT 1860
Db       1801 TGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTT 1860
Qy       1861 TCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACA 1920
Db       1861 TCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACA 1920
Qy       1921 GCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGC 1980
Db       1921 GCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGC 1980
Qy       1981 TCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTAC 2040
Db       1981 TCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTAC 2040
Qy       2041 TACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAG 2100
Db       2041 TACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAG 2100
Qy       2101 CCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTG 2160
Db       2101 CCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTG 2160
Qy       2161 GTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTG 2220
Db       2161 GTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTG 2220
Qy       2221 TTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTT 2280
Db       2221 TTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTT 2280
Qy       2281 GTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACC 2340
Db       2281 GTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACC 2340
Qy       2341 CCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCA 2400
Db       2341 CCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCA 2400
Qy       2401 TTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGA 2460
Db       2401 TTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGA 2460
Qy       2461 ATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAAT 2520
Db       2461 ATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAAT 2520
Qy       2521 GGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCC 2580
Db       2521 GGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCC 2580
```

```
Qy       2581 TTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAAC 2640
Db       2581 TTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAAC 2640
Qy       2641 ATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATG 2700
Db       2641 ATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATG 2700
Qy       2701 GAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGT 2760
Db       2701 GAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGT 2760
Qy       2761 CATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGG 2820
Db       2761 CATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGG 2820
Qy       2821 AAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGA 2880
Db       2821 AAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGA 2880
Qy       2881 AATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAAC 2940
Db       2881 AATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAAC 2940
Qy       2941 TGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAG 3000
Db       2941 TGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAG 3000
Qy       3001 AGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCT 3060
Db       3001 AGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCT 3060
Qy       3061 GGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCAC 3120
Db       3061 GGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCAC 3120
Qy       3121 GCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGG 3180
Db       3121 GCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGG 3180
Qy       3181 GACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACA 3240
Db       3181 GACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACA 3240
Qy       3241 TGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGAT 3300
Db       3241 TGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGAT 3300
Qy       3301 CGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGAC 3360
Db       3301 CGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGAC 3360
Qy       3361 ATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCA 3420
Db       3361 ATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCA 3420
Qy       3421 GCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTAC 3480
Db       3421 GCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTAC 3480
Qy       3481 AAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTT 3540
Db       3481 AAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTT 3540
Qy       3541 GTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGGTGCAAACTGAACAATGA 3597
Db       3541 GTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGGTGCAAACTGAACAATGA 3597
```

RESULT 4
AAZ52275
ID AAZ52275 standard; cDNA; 4068 BP.

```
XX
AC   AAZ52275;
XX
DT   18-JUL-2000  (first entry)
XX
DE   Human membrane attractin-1 cDNA.
XX
KW   Human; membrane attractin-1; immune response; macrophage; monocyte;
KW   T cell; immunostimulant; immunosuppressed patient; cancer;
KW   immunodeficiency syndrome; transplant; autoimmune disease; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .4068
FT                   /*tag= a
FT                   /product= "Membrane attractin-1"
XX
PN   WO200015651-A1.
XX
PD   23-MAR-2000.
XX
PF   14-SEP-1999;   99WO-US020948.
XX
PR   14-SEP-1998;   98US-0100137P.
XX
PA   (DAND ) DANA FARBER CANCER INST INC.
XX
PI   Duke-Cohan JS,  Schlossman SF;
XX
DR   WPI; 2000-271373/23.
DR   P-PSDB; AAY70690.
XX
PT   Isolated nucleic acids encoding human attractin polypeptides useful for
PT   enhancing immune responses.
XX
PS   Claim 3; Fig 10; 120pp; English.
XX
CC   The patent discloses four forms of human attractin polypeptides which
CC   enhance immune response by promoting macrophage and monocyte spreading in
CC   the presence of T cells. These include soluble attractin-1 and -2 and
CC   membrane attractin-1 and -2. These various forms of attractin are encoded
CC   by alternatively spliced mRNA molecule transcribed from a single gene.
CC   The present sequence is a cDNA encoding human membrane attractin-1.
CC   Membrane attractin differs from soluble attractin in having a
CC   transmembrane domain and a cytoplasmic domain at the C-terminal. This
CC   sequence can be used to enhance immune response in immunosuppressed
CC   patients such as those undergoing chemo- and radio-therapy treatment for
CC   cancer or those suffering from common variable immunodeficiency syndrome.
CC   The proteins may also be used to screen modulators (agonists and
CC   antagonists) of immune response which may also be used to regulate immune
CC   reactions. Attractin antibodies can be used to inhibit immune response in
CC   transplant recipients or patients afflicted with autoimmune disease
XX
SQ   Sequence 4068 BP; 1042 A; 948 C; 1027 G; 1051 T; 0 U; 0 Other;

  Query Match           98.9%;  Score 3556.4;  DB 3;  Length 4068;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches 3579;  Conservative    0;  Mismatches    1;  Indels    2;  Gaps    2;

Qy          1 ATGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 60

Qy         61 GCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGC 120

Qy        121 ACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGC 180

Qy        181 CGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATAC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATAC 240

Qy        241 AAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTC 300

Qy        301 AATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATT 360

Qy        361 TATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAG 420

Qy        421 ACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCT 480

Qy        481 GCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGC 540

Qy        541 TCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCT 600

Qy        601 GAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCT 660

Qy        661 CATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAG 720

Qy        721 GGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATAT 780

Qy        781 TCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGG 840

Qy        841 GTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTT 900

Qy        901 GCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGT 960

Qy        961 CATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACT 1020

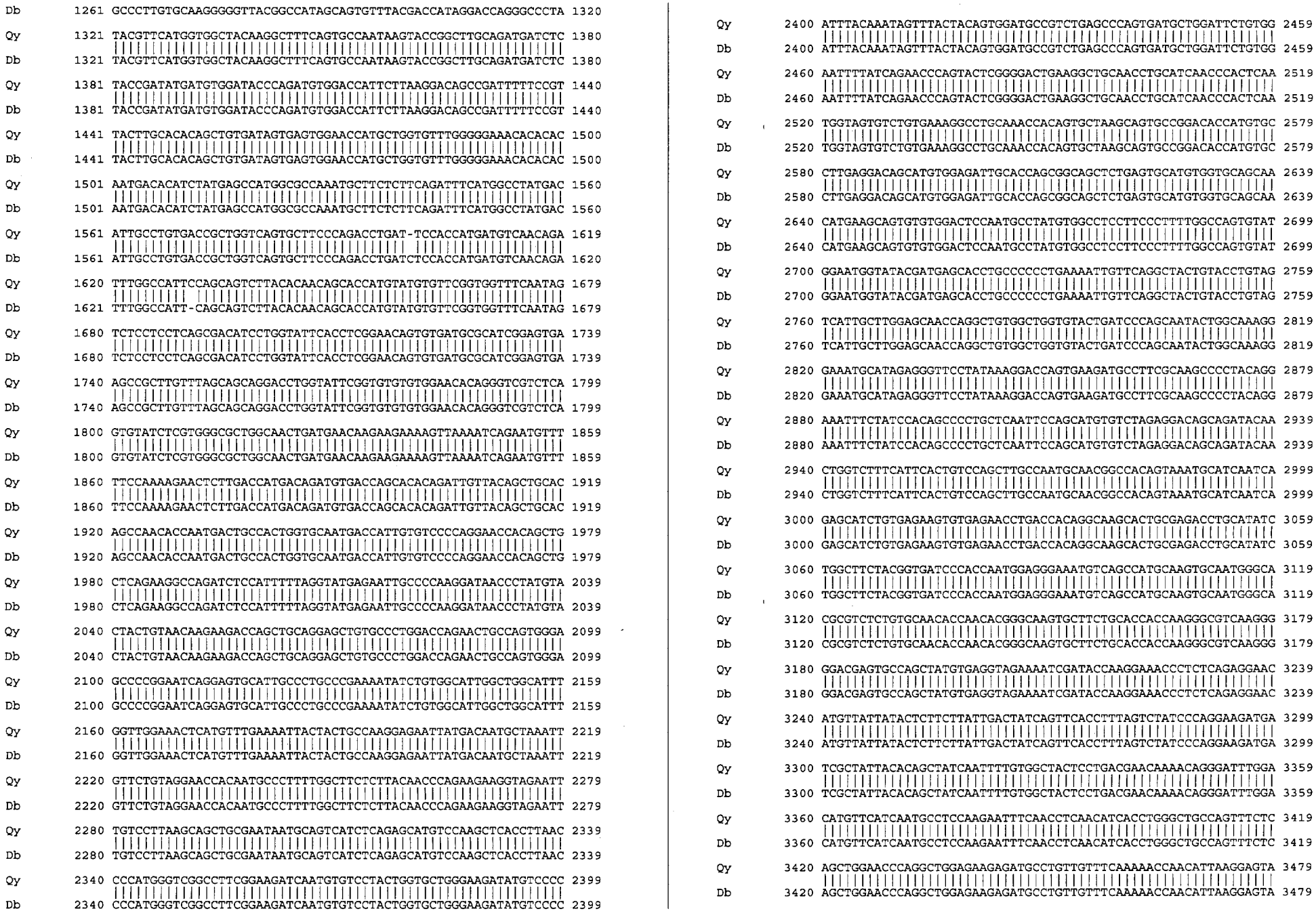
Qy       1021 GGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTG 1080

Qy       1081 ACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTG 1140

Qy       1141 AAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATA 1200

Qy       1201 AGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGT 1260

Qy       1261 GCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```



```
Qy       3480 CAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTT 3539
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3480 CAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTT 3539

Qy       3540 TGTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGGT 3581
             |||||||||||||||||||||||||||||||||||||||||||
Db       3540 TGTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGAT 3581


RESULT 5
AAZ91921
ID   AAZ91921 standard; cDNA; 4072 BP.
XX
AC   AAZ91921;
XX
DT   08-JUN-2000  (first entry)
XX
DE   Human mahogany protein coding sequence #3.
XX
KW   Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;
KW   weight regulation; cell therapy; body weight disorder; cachexia;
KW   anorexia; hyperpigmentation; increased metabolic rate disorder;
KW   hyperphagia; Antiobesity; antianorexic; anticachexic; ss.
XX
OS   Homo sapiens.
XX
PN   WO200005373-A2.
XX
PD   03-FEB-2000.
XX
PF   21-JUL-1999;   99WO-US016484.
XX
PR   21-JUL-1998;   98US-0093630P.
PR   20-OCT-1998;   98US-0104978P.
PR   05-FEB-1999;   99US-00245041.
XX
PA   (MILL-) MILLENIUM PHARM INC.
XX
PI   Moore K,  Nagle DL;
XX
DR   WPI; 2000-195103/17.
DR   P-PSDB; AAY81808.
XX
PT   New human and murine mahogany genes, useful, e.g. for diagnosis and
PT   treatment of body weight disorders.
XX
PS   Claim 1; Fig 19a; 188pp; English.
XX
CC   This sequence represents a human mahogany gene of the invention. The
CC   mahogany genes are used: (i) to produce recombinant mahogany (mg)
CC   proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming
CC   therapeutics; (iii) as a source of diagnostic probes and primers for
CC   detecting expression of mg genes or mutations, regulatory defects, in
CC   this gene, or for isolation of related sequences; and (iv) in (cell-
CC   based) gene therapy. (II) are used to raise specific antibodies (Ab); to
CC   identify other (extra)cellular products involved in weight regulation,
CC   and to screen for agents that disrupt interaction between (II) and other
CC   macromolecules. The Ab are used to detect abnormal levels (or function)
CC   of (II) (for diagnosis, prognosis or monitoring of treatment); to
CC   evaluate (II)-expressing cells intended for cell therapy, and as
CC   therapeutic mg inhibitors. Cells that express the mg gene (or contain the
CC   mg polypeptide) are used to identify agents (A) that modulate mg
CC   activity. (A) are potentially useful for the treatment of body weight
CC   disorders, particularly obesity, cachexia or anorexia, or other
CC   conditions associated with the mg gene such as hyperpigmentation,
CC   hyperphagia and disorders that result in increased metabolic rate
XX
SQ   Sequence 4072 BP; 1087 A; 940 C; 1045 G; 1000 T; 0 U; 0 Other;

  Query Match             97.1%;  Score 3494.4;  DB 3;  Length 4072;
  Best Local Similarity   98.8%;  Pred. No. 0;
  Matches 3541;  Conservative    0;  Mismatches   41;  Indels    2;  Gaps    2;

Qy         15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74
              || |||     |  ||   | ||  |  |  | | |    |||| ||   || || ||
Db        234 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAA 293

Qy         75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134
              |    |  |  ||  ||||||||||||||||||||||||||||||||||||||||||||
Db        294 GGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 353

Qy        135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        354 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 413

Qy        195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 473

Qy        255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 533

Qy        315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 593

Qy        375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 653

Qy        435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        654 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 713

Qy        495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        714 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 773

Qy        555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        774 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 833

Qy        615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        834 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 893

Qy        675 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        894 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 953

Qy        735 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        954 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 1013

Qy        795 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1014 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 1073

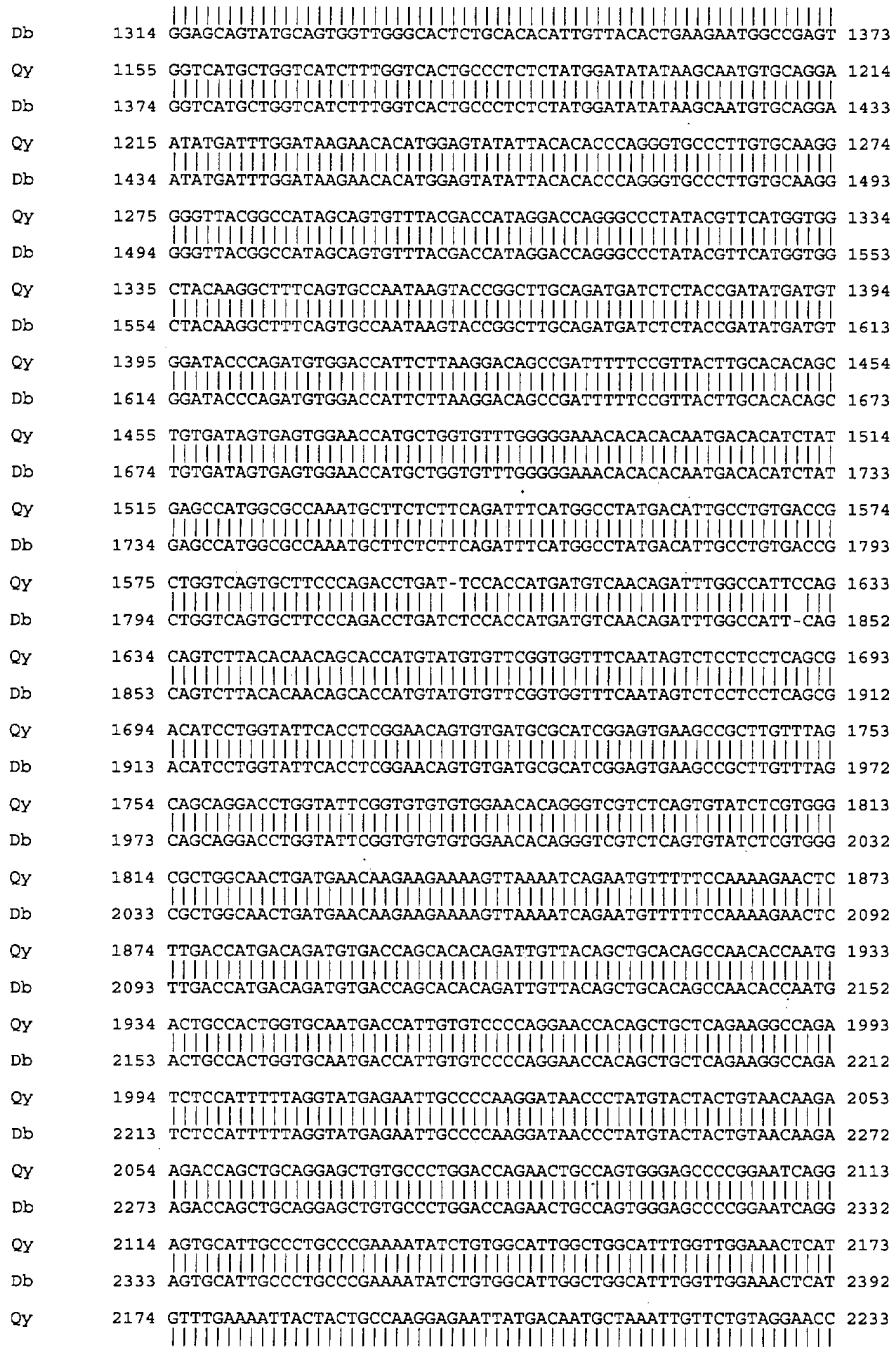
Qy        855 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1074 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 1133

Qy        915 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1134 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 1193

Qy        975 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1194 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1253

Qy       1035 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1254 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1313

Qy       1095 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1154
```


Db 1314 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1373
Qy 1155 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1214
Db 1374 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1433
Qy 1215 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1274
Db 1434 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1493
Qy 1275 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1334
Db 1494 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1553
Qy 1335 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1394
Db 1554 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1613
Qy 1395 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1454
Db 1614 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1673
Qy 1455 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1514
Db 1674 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1733
Qy 1515 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1574
Db 1734 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1793
Qy 1575 CTGGTCAGTGCTTCCCAGACCTGAT-TCCACCATGATGTCAACAGATTTGGCCATTCCAG 1633
Db 1794 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATT-CAG 1852
Qy 1634 CAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCG 1693
Db 1853 CAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCG 1912
Qy 1694 ACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAG 1753
Db 1913 ACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAG 1972
Qy 1754 CAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGG 1813
Db 1973 CAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGG 2032
Qy 1814 CGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTC 1873
Db 2033 CGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTC 2092
Qy 1874 TTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATG 1933
Db 2093 TTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATG 2152
Qy 1934 ACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGA 1993
Db 2153 ACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGA 2212
Qy 1994 TCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGA 2053
Db 2213 TCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGA 2272
Qy 2054 AGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGG 2113
Db 2273 AGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGG 2332
Qy 2114 AGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCAT 2173
Db 2333 AGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCAT 2392
Qy 2174 GTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACC 2233
Db 2393 GTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACC 2452
Qy 2234 ACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGC 2293
Db 2453 ACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGC 2512
Qy 2294 TGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCC 2353
Db 2513 TGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCC 2572
Qy 2354 TTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTT 2413
Db 2573 TTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTT 2632
Qy 2414 TACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAAC 2473
Db 2633 TACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAAC 2692
Qy 2474 CCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTG 2533
Db 2693 CCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTG 2752
Qy 2534 AAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCAT 2593
Db 2753 AAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCAT 2812
Qy 2594 GTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTG 2653
Db 2813 GTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTG 2872
Qy 2654 TGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGA 2713
Db 2873 TGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGA 2932
Qy 2714 TGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGC 2773
Db 2933 TGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGC 2992
Qy 2774 AACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGG 2833
Db 2993 AACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGG 3052
Qy 2834 GTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCAC 2893
Db 3053 GTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCAC 3112
Qy 2894 AGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTC 2953
Db 3113 AGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTC 3172
Qy 2954 ACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGA 3013
Db 3173 ACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGA 3232
Qy 3014 AGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTG 3073
Db 3233 AGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTG 3292
Qy 3074 ATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCA 3133
Db 3293 ATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCA 3352
Qy 3134 ACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGC 3193
Db 3353 ACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGC 3412
Qy 3194 TATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTC 3253
Db 3413 TATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTC 3472
Qy 3254 TTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAG 3313
Db 3473 TTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAG 3532

```
Qy       3314 CTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATG 3373
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3533 CTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATG 3592

Qy       3374 CCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGG 3433
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3593 CCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGG 3652

Qy       3434 CTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCT 3493
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3653 CTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCT 3712

Qy       3494 CTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTA 3553
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3713 CTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTA 3772

Qy       3554 ATTTCACCTGGCCCATCAAAATTCAGGTGCAAACTGAACAATGA 3597
             |||||||||||||||||||||||||||||||||||||||||||
Db       3773 ATTTCACCTGGCCCATCAAAATTCAGGTGCAAACTGAACAATGA 3816
```

RESULT 6
AAZ52277
ID AAZ52277 standard; cDNA; 3819 BP.
XX
AC AAZ52277;
XX
DT 18-JUL-2000 (first entry)
XX
DE Human soluble attractin-2 cDNA.
XX
KW Human; soluble attractin-2; immune response; macrophage; monocyte;
KW T cell; immunostimulant; immunosuppressed patient; cancer;
KW immunodeficiency syndrome; transplant; autoimmune disease; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1. .3819
FT /*tag= a
FT /product= "Soluble attractin-2"
XX
PN WO200015651-A1.
XX
PD 23-MAR-2000.
XX
PF 14-SEP-1999; 99WO-US020948.
XX
PR 14-SEP-1998; 98US-0100137P.
XX
PA (DAND) DANA FARBER CANCER INST INC.
XX
PI Duke-Cohan JS, Schlossman SF;
XX
DR WPI; 2000-271373/23.
DR P-PSDB; AAY70692.
XX
PT Isolated nucleic acids encoding human attractin polypeptides useful for
PT enhancing immune responses.
XX
PS Claim 3; Fig 12; 120pp; English.
XX
CC The patent discloses four forms of human attractin polypeptides which
CC enhance immune response by promoting macrophage and monocyte spreading in
CC the presence of T cells. These include soluble attractin-1 and -2 and
CC membrane attractin-1 and -2. These various forms of attractin are encoded
CC by alternatively spliced mRNA molecule transcribed from a single gene.
CC The present sequence is a cDNA encoding human soluble attractin-2.
CC Attractin-2 differs from attractin-1 in having a 74 amino acid insert in
CC the N-terminal. This sequence can be used to enhance immune response in
CC immunosuppressed patients such as those undergoing chemo- and radio-
CC therapy treatment for cancer or those suffering from common variable
CC immunodeficiency syndrome. The proteins may also be used to screen
CC modulators (agonists and antagonists) of immune response which may also
CC be used to regulate immune reactions. Attractin antibodies can be used to
CC inhibit immune response in transplant recipients or patients afflicted
CC with autoimmune disease
XX
SQ Sequence 3819 BP; 969 A; 899 C; 991 G; 960 T; 0 U; 0 Other;

Query Match 96.9%; Score 3487; DB 3; Length 3819;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 3524; Conservative 0; Mismatches 25; Indels 2; Gaps 2;

```
Qy         48 GGCGGCGACGGCAGCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGTGTCAACGGCGGTCG 107
              |||  |  || ||  || || || |  |  | || ||||||||||||||||||||||||
Db        270 GTCGGGCTCAGCCGCAGCCGAGGCCAAGGAATGTGACCGGCCCTGTGTCAACGGCGGTCG 329

Qy        108 CTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCA 167
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        330 CTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCA 389

Qy        168 GCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGG 227
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        390 GCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGG 449

Qy        228 AAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAAT 287
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        450 AAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAAT 509

Qy        288 GAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGA 347
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        510 GAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGA 569

Qy        348 TGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAG 407
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        570 TGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAG 629

Qy        408 AGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTT 467
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        630 AGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTT 689

Qy        468 TTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTG 527
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        690 TTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTG 749

Qy        528 TCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGA 587
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        750 TCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGAAACTGTTGA 809

Qy        588 ATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA 647
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        810 ATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA 869

Qy        648 CTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTT 707
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        870 CTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTT 929

Qy        708 CTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGAC 767
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        930 CTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGAC 989

Qy        768 TCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGG 827
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        990 TCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGG 1049

Qy        828 AAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCT 887
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1050 AAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCT 1109

Qy        888 AGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGT 947
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1110 AGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGT 1169

Qy        948 TGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAA 1007
```

```
Db      1170 TGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAA 1229
Qy      1008 AATTGATTCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTC 1067
Db      1230 AATTGATCCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTC 1289
Qy      1068 ATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACA 1127
Db      1290 ATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACA 1349
Qy      1128 CATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCT 1187
Db      1350 CATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCT 1409
Qy      1188 CTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATT 1247
Db      1410 CTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATT 1469
Qy      1248 ACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAG 1307
Db      1470 ACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAG 1529
Qy      1308 GACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCT 1367
Db      1530 GACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCT 1589
Qy      1368 TGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAG 1427
Db      1590 TGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAG 1649
Qy      1428 CCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGG 1487
Db      1650 CCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGG 1709
Qy      1488 GGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTT 1547
Db      1710 GGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTT 1769
Qy      1548 CATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGAT-TCCACCA 1606
Db      1770 CATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCA 1829
Qy      1607 TGATGTCAACAGATTTGGCCATTCCAGCAGTCTTACACAACAGCACCATGTATGTGTTCG 1666
Db      1830 TGATGTCAACAGATTTGGCCATT-CAGCAGTCTTACACAACAGCACCATGTATGTGTTCG 1888
Qy      1667 GTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATG 1726
Db      1889 GTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATG 1948
Qy      1727 CGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACA 1786
Db      1949 CGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACA 2008
Qy      1787 CAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAA 1846
Db      2009 CAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAA 2068
Qy      1847 AATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATT 1906
Db      2069 AATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATT 2128
Qy      1907 GTTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCA 1966
Db      2129 GTTACAGCTGTACAGCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCA 2188
Qy      1967 GGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGG 2026
Db      2189 GGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGG 2248
Qy      2027 ATAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGA 2086
```

```
Db      2249 ATAACCCCATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGA 2308
Qy      2087 ACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCA 2146
Db      2309 ACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCA 2368
Qy      2147 TTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATG 2206
Db      2369 TTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATG 2428
Qy      2207 ACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGA 2266
Db      2429 ACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGA 2488
Qy      2267 AGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCA 2326
Db      2489 AGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCA 2548
Qy      2327 AGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGG 2386
Db      2549 AGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGG 2608
Qy      2387 AAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATG 2446
Db      2609 AAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATG 2668
Qy      2447 CTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCA 2506
Db      2669 CTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCA 2728
Qy      2507 TCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCC 2566
Db      2729 TCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCC 2788
Qy      2567 GGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCA 2626
Db      2789 GGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCA 2848
Qy      2627 TGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTT 2686
Db      2849 TGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTT 2908
Qy      2687 TTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCT 2746
Db      2909 TTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCT 2968
Qy      2747 ACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCA 2806
Db      2969 ACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCA 3028
Qy      2807 ATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGC 2866
Db      3029 ATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGC 3088
Qy      2867 AAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGG 2926
Db      3089 AAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGG 3148
Qy      2927 ACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTA 2986
Db      3149 ACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTA 3208
Qy      2987 AATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCG 3046
Db      3209 AATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCG 3268
Qy      3047 AGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCA 3106
Db      3269 AGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCA 3328
Qy      3107 AGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCA 3166
Db      3329 AGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCA 3388
```

```
Qy       3167 AGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACC 3226
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3389 AGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACC 3448

Qy       3227 CTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTAT 3286
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3449 CTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTAT 3508

Qy       3287 CCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAA 3346
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3509 CCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAA 3568

Qy       3347 ACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGG 3406
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3569 ACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGG 3628

Qy       3407 CTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCA 3466
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3629 CTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCA 3688

Qy       3467 ACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAA 3526
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3689 ACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAA 3748

Qy       3527 ATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGGTGCAAA 3586
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3749 ATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGGTGCAAA 3808

Qy       3587 CTGAACAATGA 3597
             ||||||||||||
Db       3809 CTGAACAATGA 3819
```

RESULT 7
AAZ91920
ID AAZ91920 standard; cDNA; 8590 BP.
XX
AC AAZ91920;
XX
DT 08-JUN-2000 (first entry)
XX
DE Human mahogany protein coding sequence #2.
XX
KW Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;
KW weight regulation; cell therapy; body weight disorder; cachexia;
KW anorexia; hyperpigmentation; increased metabolic rate disorder;
KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.
XX
OS Homo sapiens.
XX
PN WO200005373-A2.
XX
PD 03-FEB-2000.
XX
PF 21-JUL-1999; 99WO-US016484.
XX
PR 21-JUL-1998; 98US-0093630P.
PR 20-OCT-1998; 98US-0104978P.
PR 05-FEB-1999; 99US-00245041.
XX
PA (MILL-) MILLENIUM PHARM INC.
XX
PI Moore K, Nagle DL;
XX
DR WPI; 2000-195103/17.
DR P-PSDB; AAY81807.
XX
PT New human and murine mahogany genes, useful, e.g. for diagnosis and
PT treatment of body weight disorders.
XX
PS Claim 1; Fig 18a; 188pp; English.
XX
CC This sequence represents a human mahogany gene of the invention. The
CC mahogany genes are used: (i) to produce recombinant mahogany (mg)
CC proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming
CC therapeutics; (iii) as a source of diagnostic probes and primers for
CC detecting expression of mg genes or mutations, regulatory defects, in
CC this gene, or for isolation of related sequences; and (iv) in (cell-
CC based) gene therapy. (II) are used to raise specific antibodies (Ab); to
CC identify other (extra)cellular products involved in weight regulation,
CC and to screen for agents that disrupt interaction between (II) and other
CC macromolecules. The Ab are used to detect abnormal levels (or function)
CC of (II) (for diagnosis, prognosis or monitoring of treatment); to
CC evaluate (II)-expressing cells intended for cell therapy, and as
CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the
CC mg polypeptide) are used to identify agents (A) that modulate mg
CC activity. (A) are potentially useful for the treatment of body weight
CC disorders, particularly obesity, cachexia or anorexia, or other
CC conditions associated with the mg gene such as hyperpigmentation,
CC hyperphagia and disorders that result in increased metabolic rate
XX
SQ Sequence 8590 BP; 2275 A; 1990 C; 2061 G; 2264 T; 0 U; 0 Other;

Query Match 96.7%; Score 3476.8; DB 3; Length 8590;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 3524; Conservative 0; Mismatches 42; Indels 2; Gaps 2;

```
Qy         15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74
              || |||   |  ||   | ||  |  |  |  | | |   |||| ||   || || ||
Db        234 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAA 293

Qy         75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134
              |   | |  | ||  ||||||||||||||||||||||||||||||||||||||||||||
Db        294 GGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 353

Qy        135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGGCCGCTTCAGACTAAC 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        354 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 413

Qy        195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 473

Qy        255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 533

Qy        315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 593

Qy        375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 653

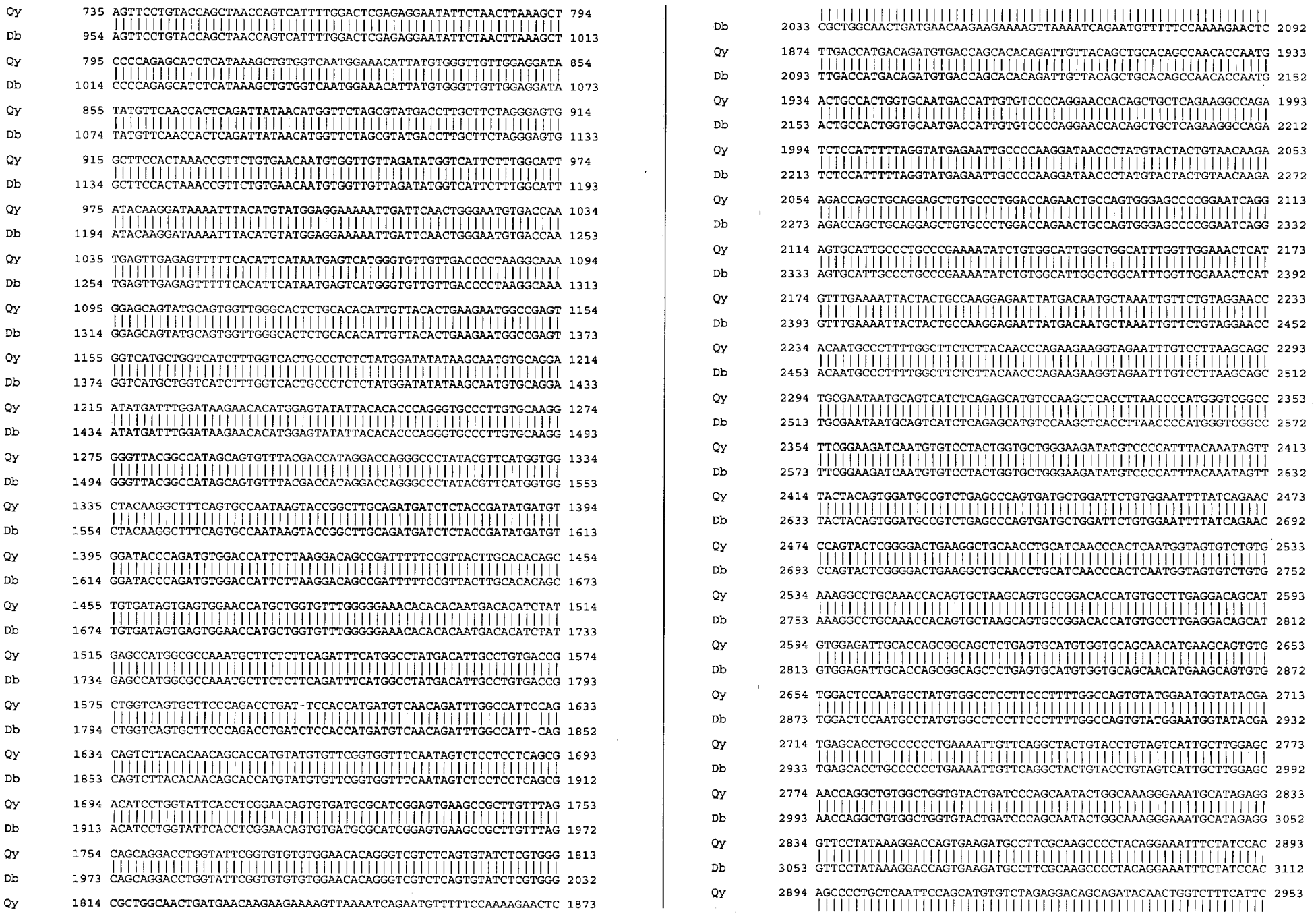
Qy        435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        654 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 713

Qy        495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        714 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 773

Qy        555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        774 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 833

Qy        615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        834 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 893

Qy        675 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        894 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 953
```



```
Db       3113 AGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTC 3172

Qy       2954 ACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGA 3013
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3173 ACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGA 3232

Qy       3014 AGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTG 3073
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3233 AGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTG 3292

Qy       3074 ATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCA 3133
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3293 ATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCA 3352

Qy       3134 ACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGC 3193
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3353 ACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGC 3412

Qy       3194 TATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTC 3253
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3413 TATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTC 3472

Qy       3254 TTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAG 3313
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3473 TTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAG 3532

Qy       3314 CTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATG 3373
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3533 CTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATG 3592

Qy       3374 CCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGG 3433
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3593 CCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGG 3652

Qy       3434 CTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCT 3493
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3653 CTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCT 3712

Qy       3494 CTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTA 3553
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3713 CTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTA 3772

Qy       3554 ATTTCACCTGGCCCATCAAAATTCAGGT 3581
             |||||||||||||||||||||||||| |
Db       3773 ATTTCACCTGGCCCATCAAAATTCAGAT 3800
```

```
RESULT 8
AAZ52276
ID   AAZ52276 standard; cDNA; 4290 BP.
XX
AC   AAZ52276;
XX
DT   18-JUL-2000  (first entry)
XX
DE   Human membrane attractin-2 cDNA.
XX
KW   Human; membrane attractin-2; immune response; macrophage; monocyte;
KW   T cell; immunostimulant; immunosuppressed patient; cancer;
KW   immunodeficiency syndrome; transplant; autoimmune disease; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .4290
FT                   /*tag=  a
FT                   /product= "Membrane attractin-2"
XX
PN   WO200015651-A1.
XX
PD   23-MAR-2000.
XX
PF   14-SEP-1999;   99WO-US020948.
XX
PR   14-SEP-1998;   98US-0100137P.
XX
PA   (DAND ) DANA FARBER CANCER INST INC.
XX
PI   Duke-Cohan JS,  Schlossman SF;
XX
DR   WPI; 2000-271373/23.
DR   P-PSDB; AAY70691.
XX
PT   Isolated nucleic acids encoding human attractin polypeptides useful for
PT   enhancing immune responses.
XX
PS   Claim 3; Fig 14; 120pp; English.
XX
CC   The patent discloses four forms of human attractin polypeptides which
CC   enhance immune response by promoting macrophage and monocyte spreading in
CC   the presence of T cells. These include soluble attractin-1 and -2 and
CC   membrane attractin-1 and -2. These various forms of attractin are encoded
CC   by alternatively spliced mRNA molecule transcribed from a single gene.
CC   The present sequence is a cDNA encoding human membrane attractin-2.
CC   Membrane attractin differs from soluble attractin in having a
CC   transmembrane domain and a cytoplasmic domain at the C-terminal.
CC   Attractin-2 differs from attractin-1 in having a 74 amino acid insert in
CC   the N-terminal. This sequence can be used to enhance immune response in
CC   immunosuppressed patients such as those undergoing chemo- and radio-
CC   therapy treatment for cancer or those suffering from common variable
CC   immunodeficiency syndrome. The proteins may also be used to screen
CC   modulators (agonists and antagonists) of immune response which may also
CC   be used to regulate immune reactions. Attractin antibodies can be used to
CC   inhibit immune response in transplant recipients or patients afflicted
CC   with autoimmune disease
XX
SQ   Sequence 4290 BP; 1060 A; 1025 C; 1120 G; 1085 T; 0 U; 0 Other;
```

```
  Query Match             96.5%;  Score 3469.4;  DB 3;  Length 4290;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 3507;  Conservative    0;  Mismatches   26;  Indels    2;  Gaps    2;

Qy         48 GGCGGCGACGGCAGCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGTGTCAACGGCGGTCG 107
              |||    | || ||   || || || |    | |  | || ||||||||||||||||||
Db        270 GTCGGGCTCAGCCGCAGCCGAGGCCAAGGAATGTGACCGGCCCTGTGTCAACGGCGGTCG 329

Qy        108 CTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCA 167
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        330 CTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCA 389

Qy        168 GCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGG 227
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        390 GCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGG 449

Qy        228 AAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAAT 287
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        450 AAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAAT 509

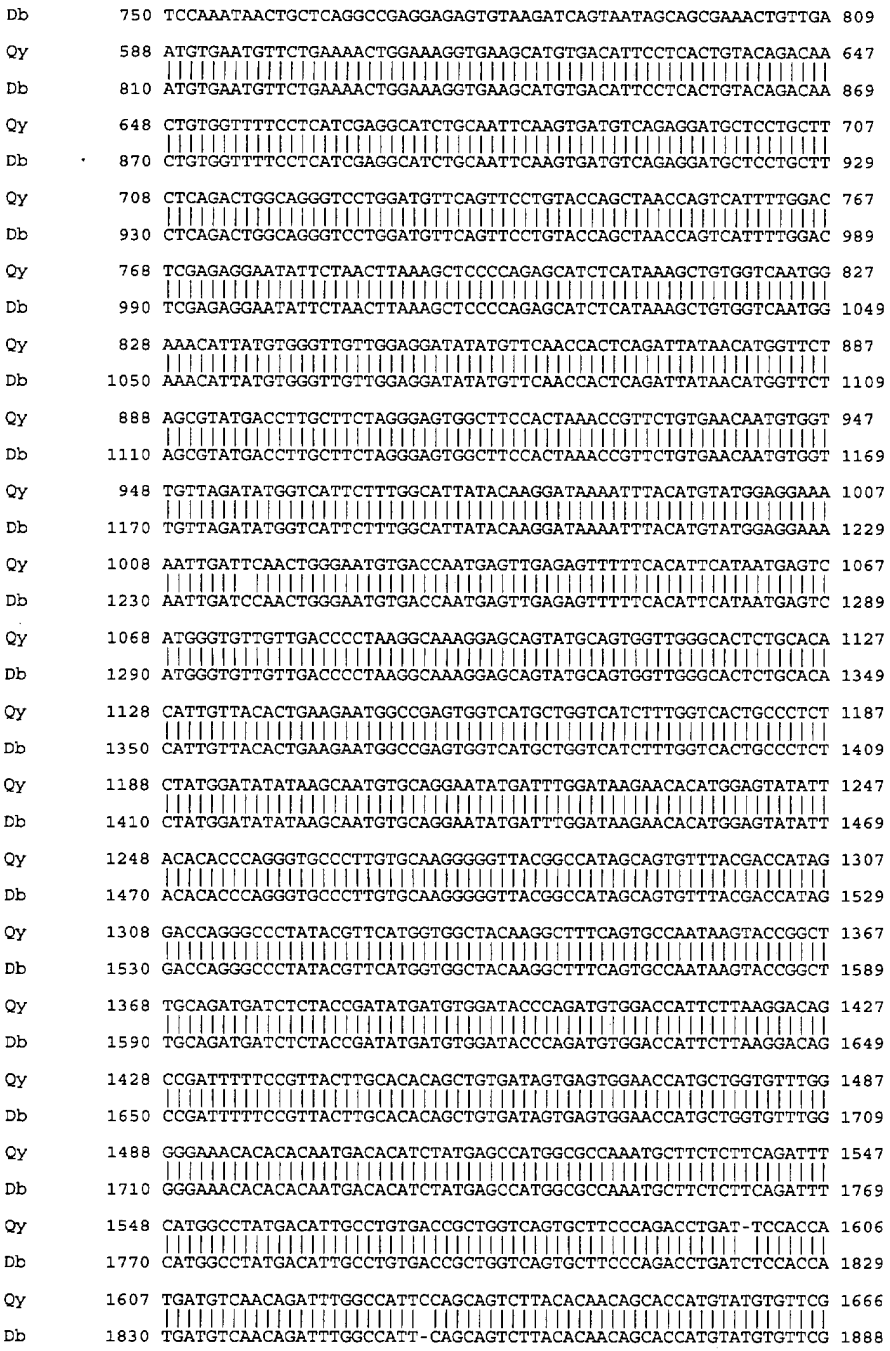
Qy        288 GAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGA 347
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        510 GAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGA 569

Qy        348 TGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAG 407
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        570 TGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAG 629

Qy        408 AGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTT 467
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        630 AGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTT 689

Qy        468 TTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTG 527
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        690 TTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTG 749

Qy        528 TCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGA 587
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
```


Db 750 TCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGAAACTGTTGA 809
Qy 588 ATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA 647
Db 810 ATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA 869
Qy 648 CTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTT 707
Db 870 CTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTT 929
Qy 708 CTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGAC 767
Db 930 CTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGAC 989
Qy 768 TCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGG 827
Db 990 TCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGG 1049
Qy 828 AAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCT 887
Db 1050 AAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCT 1109
Qy 888 AGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGT 947
Db 1110 AGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGT 1169
Qy 948 TGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAA 1007
Db 1170 TGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAA 1229
Qy 1008 AATTGATTCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTC 1067
Db 1230 AATTGATCCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTC 1289
Qy 1068 ATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACA 1127
Db 1290 ATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACA 1349
Qy 1128 CATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCT 1187
Db 1350 CATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCT 1409
Qy 1188 CTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATT 1247
Db 1410 CTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATT 1469
Qy 1248 ACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAG 1307
Db 1470 ACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAG 1529
Qy 1308 GACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCT 1367
Db 1530 GACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCT 1589
Qy 1368 TGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAG 1427
Db 1590 TGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAG 1649
Qy 1428 CCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGG 1487
Db 1650 CCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGG 1709
Qy 1488 GGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTT 1547
Db 1710 GGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTT 1769
Qy 1548 CATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGAT-TCCACCA 1606
Db 1770 CATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCA 1829
Qy 1607 TGATGTCAACAGATTTGGCCATTCCAGCAGTCTTACACAACAGCACCATGTATGTGTTCG 1666
Db 1830 TGATGTCAACAGATTTGGCCATT-CAGCAGTCTTACACAACAGCACCATGTATGTGTTCG 1888
Qy 1667 GTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATG 1726
Db 1889 GTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATG 1948
Qy 1727 CGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACA 1786
Db 1949 CGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACA 2008
Qy 1787 CAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAA 1846
Db 2009 CAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAA 2068
Qy 1847 AATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATT 1906
Db 2069 AATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATT 2128
Qy 1907 GTTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCA 1966
Db 2129 GTTACAGCTGTACAGCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCA 2188
Qy 1967 GGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGG 2026
Db 2189 GGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGG 2248
Qy 2027 ATAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGA 2086
Db 2249 ATAACCCCATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGA 2308
Qy 2087 ACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCA 2146
Db 2309 ACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCA 2368
Qy 2147 TTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATG 2206
Db 2369 TTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATG 2428
Qy 2207 ACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGA 2266
Db 2429 ACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGA 2488
Qy 2267 AGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCA 2326
Db 2489 AGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCA 2548
Qy 2327 AGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGG 2386
Db 2549 AGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGG 2608
Qy 2387 AAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATG 2446
Db 2609 AAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATG 2668
Qy 2447 CTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCA 2506
Db 2669 CTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCA 2728
Qy 2507 TCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCC 2566
Db 2729 TCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCC 2788
Qy 2567 GGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCA 2626
Db 2789 GGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCA 2848
Qy 2627 TGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTT 2686
Db 2849 TGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTT 2908
Qy 2687 TTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCT 2746
Db 2909 TTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCT 2968

```
Qy       2747 ACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCA 2806
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2969 ACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCA 3028

Qy       2807 ATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGC 2866
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3029 ATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGC 3088

Qy       2867 AAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGG 2926
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3089 AAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGG 3148

Qy       2927 ACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTA 2986
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3149 ACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTA 3208

Qy       2987 AATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCG 3046
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3209 AATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCG 3268

Qy       3047 AGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCA 3106
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3269 AGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCA 3328

Qy       3107 AGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCA 3166
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3329 AGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCA 3388

Qy       3167 AGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACC 3226
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3389 AGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACC 3448

Qy       3227 CTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTAT 3286
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3449 CTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTAT 3508

Qy       3287 CCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAA 3346
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3509 CCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAA 3568

Qy       3347 ACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGG 3406
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3569 ACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGG 3628

Qy       3407 CTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCA 3466
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3629 CTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCA 3688

Qy       3467 ACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAA 3526
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3689 ACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAA 3748

Qy       3527 ATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGGT 3581
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||  |
Db       3749 ATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGAT 3803
```

```
RESULT 9
AAS72659
ID   AAS72659 standard; cDNA; 8561 BP.
XX
AC   AAS72659;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #8463.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG08472.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 8463; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 8561 BP; 2246 A; 1983 C; 2065 G; 2267 T; 0 U; 0 Other;

  Query Match           96.5%;  Score 3469.4;  DB 5;  Length 8561;
  Best Local Similarity  99.2%;  Pred. No. 0;
  Matches 3507;  Conservative    0;  Mismatches   26;  Indels    2;  Gaps    2;

Qy         48 GGCGGCGACGGCAGCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGTGTCAACGGCGGTCG 107
              | |||   | || ||   || || || |    | |  | || ||||||||||||||||||
Db        270 GTCGGGCTCAGCCGCAGCCGAGGCCAAGGAATGTGACCGGCCCTGTGTCAACGGCGGTCG 329

Qy        108 CTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCA 167
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        330 CTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCA 389

Qy        168 GCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGG 227
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        390 GCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGG 449

Qy        228 AAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAAT 287
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        450 AAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAAT 509

Qy        288 GAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGA 347
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        510 GAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGA 569

Qy        348 TGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAG 407
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        570 TGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAG 629
```

```
Qy       408 AGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTT 467
Db       630 AGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTT 689

Qy       468 TTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTG 527
Db       690 TTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTG 749

Qy       528 TCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGA 587
Db       750 TCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGAAACTGTTGA 809

Qy       588 ATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA 647
Db       810 ATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA 869

Qy       648 CTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTT 707
Db       870 CTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTT 929

Qy       708 CTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGAC 767
Db       930 CTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGAC 989

Qy       768 TCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGG 827
Db       990 TCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGG 1049

Qy       828 AAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCT 887
Db      1050 AAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCT 1109

Qy       888 AGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGT 947
Db      1110 AGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGT 1169

Qy       948 TGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAA 1007
Db      1170 TGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAA 1229

Qy      1008 AATTGATTCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTC 1067
Db      1230 AATTGATCCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTC 1289

Qy      1068 ATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACA 1127
Db      1290 ATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACA 1349

Qy      1128 CATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCT 1187
Db      1350 CATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCT 1409

Qy      1188 CTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATT 1247
Db      1410 CTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATT 1469

Qy      1248 ACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAG 1307
Db      1470 ACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAG 1529

Qy      1308 GACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCT 1367
Db      1530 GACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCT 1589

Qy      1368 TGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAG 1427
Db      1590 TGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAG 1649

Qy      1428 CCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGG 1487
Db      1650 CCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGG 1709

Qy      1488 GGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTT 1547
Db      1710 GGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTT 1769

Qy      1548 CATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGAT-TCCACCA 1606
Db      1770 CATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCA 1829

Qy      1607 TGATGTCAACAGATTTGGCCATTCCAGCAGTCTTACACAACAGCACCATGTATGTGTTCG 1666
Db      1830 TGATGTCAACAGATTTGGCCATT-CAGCAGTCTTACACAACAGCACCATGTATGTGTTCG 1888

Qy      1667 GTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATG 1726
Db      1889 GTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATG 1948

Qy      1727 CGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACA 1786
Db      1949 CGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACA 2008

Qy      1787 CAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAA 1846
Db      2009 CAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAA 2068

Qy      1847 AATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATT 1906
Db      2069 AATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATT 2128

Qy      1907 GTTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCA 1966
Db      2129 GTTACAGCTGTACAGCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCA 2188

Qy      1967 GGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGG 2026
Db      2189 GGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGG 2248

Qy      2027 ATAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGA 2086
Db      2249 ATAACCCCATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGA 2308

Qy      2087 ACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCA 2146
Db      2309 ACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCA 2368

Qy      2147 TTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATG 2206
Db      2369 TTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATG 2428

Qy      2207 ACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGA 2266
Db      2429 ACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGA 2488

Qy      2267 AGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCA 2326
Db      2489 AGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCA 2548

Qy      2327 AGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGG 2386
Db      2549 AGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGG 2608

Qy      2387 AAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATG 2446
Db      2609 AAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATG 2668

Qy      2447 CTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCA 2506
Db      2669 CTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCA 2728

Qy      2507 TCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCC 2566
Db      2729 TCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCC 2788

Qy      2567 GGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCA 2626
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2789 GGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCA 2848

Qy       2627 TGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTT 2686
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2849 TGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTT 2908

Qy       2687 TTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCT 2746
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2909 TTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCT 2968

Qy       2747 ACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCA 2806
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2969 ACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCA 3028

Qy       2807 ATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGC 2866
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3029 ATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGC 3088

Qy       2867 AAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGG 2926
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3089 AAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGG 3148

Qy       2927 ACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTA 2986
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3149 ACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTA 3208

Qy       2987 AATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCG 3046
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3209 AATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCG 3268

Qy       3047 AGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCA 3106
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3269 AGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCA 3328

Qy       3107 AGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCA 3166
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3329 AGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCA 3388

Qy       3167 AGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACC 3226
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3389 AGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACC 3448

Qy       3227 CTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTAT 3286
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3449 CTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTAT 3508

Qy       3287 CCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAA 3346
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3509 CCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAA 3568

Qy       3347 ACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGG 3406
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3569 ACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGG 3628

Qy       3407 CTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCA 3466
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3629 CTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCA 3688

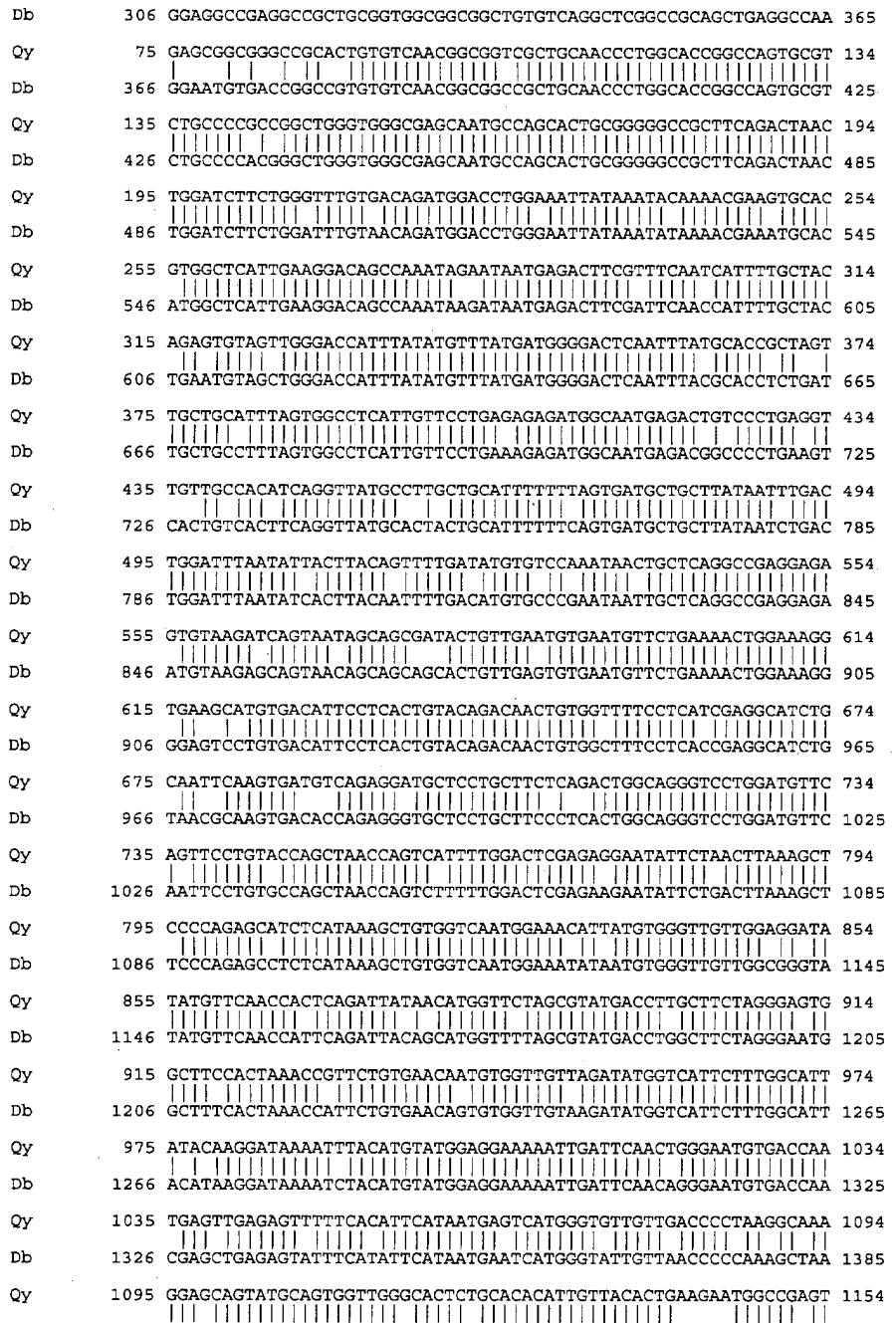
Qy       3467 ACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAA 3526
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3689 ACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAA 3748

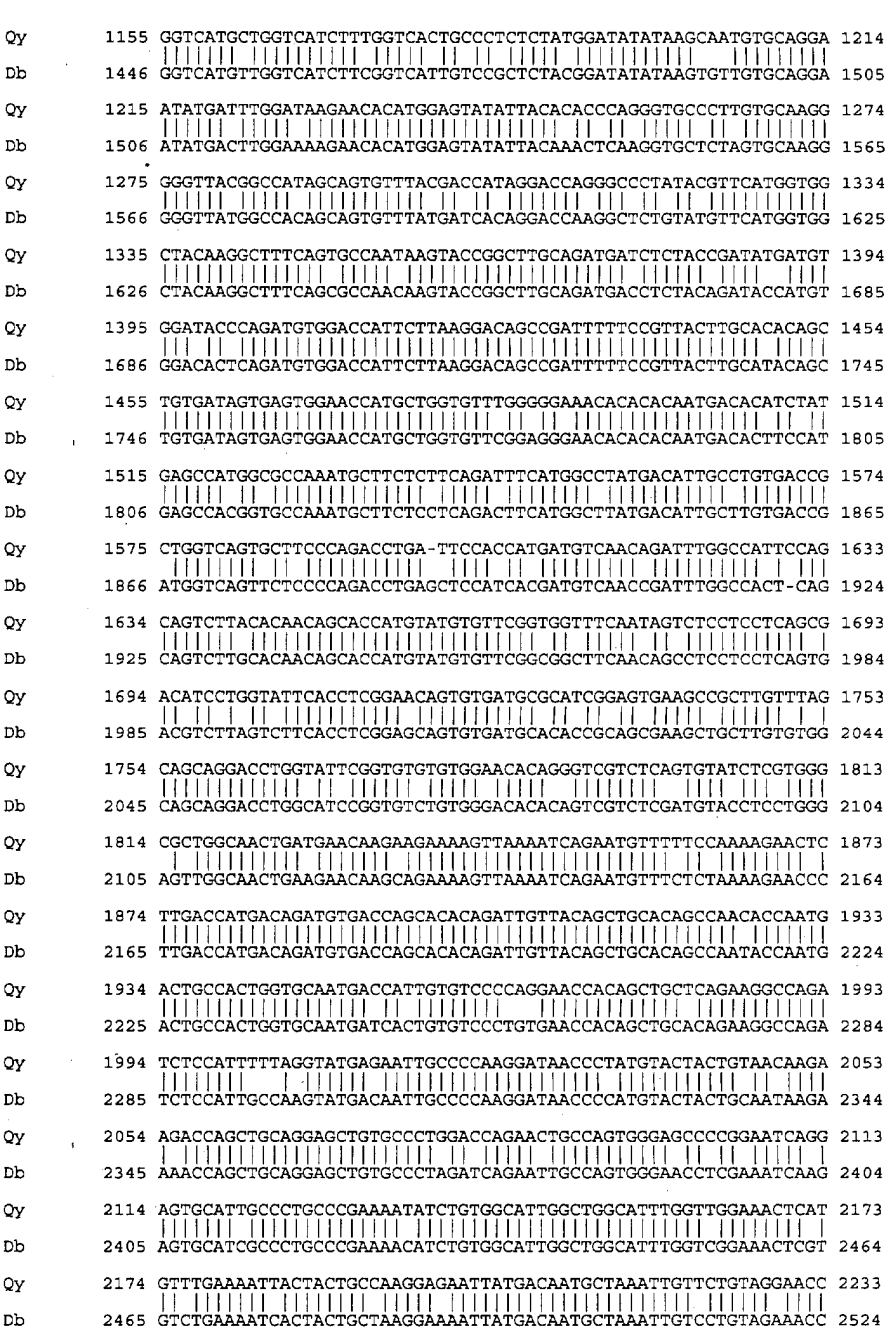
Qy       3527 ATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGGT 3581
                ||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       3749 ATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGAT 3803


RESULT 10
ADB53249
ID   ADB53249 standard; DNA; 8739 BP.
XX
AC   ADB53249;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3791.
XX
KW   toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KW   toxicity marker; toxicity progression; drug screening;
KW   primary rat hepatocyte toxicity modelling; gene; ds.
XX
OS   Rattus norvegicus.
XX
PN   WO2003065993-A2.
XX
PD   14-AUG-2003.
XX
PF   04-FEB-2003; 2003WO-US003482.
XX
PR   04-FEB-2002; 2002US-0353171P.
PR   13-MAR-2002; 2002US-0363534P.
PR   08-APR-2002; 2002US-0370248P.
PR   10-APR-2002; 2002US-0371134P.
PR   10-APR-2002; 2002US-0371135P.
PR   10-APR-2002; 2002US-0371150P.
PR   11-APR-2002; 2002US-0371413P.
PR   19-APR-2002; 2002US-0373601P.
PR   19-APR-2002; 2002US-0373602P.
PR   22-APR-2002; 2002US-0374139P.
PR   08-MAY-2002; 2002US-0378370P.
PR   09-MAY-2002; 2002US-0378652P.
PR   09-MAY-2002; 2002US-0378653P.
PR   09-MAY-2002; 2002US-0378665P.
PR   09-JUL-2002; 2002US-0394230P.
PR   09-JUL-2002; 2002US-0394253P.
PR   04-SEP-2002; 2002US-0407688P.
PR   28-JAN-2003; 2003US-0442900P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Mendrick D,  Porter M,  Johnson K,  Higgs B,  Castle A,  Orr M;
PI   Elashoff M;
XX
DR   WPI; 2003-731472/69.
XX
PT   Determining if a compound induces a toxic effect on a tissue or cell, for
PT   identifying hepatotoxic compounds, comprises comparing a gene expression
PT   profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT   mean values.
XX
PS   Claim 44; SEQ ID NO 3791; 874pp; English.
XX
CC   The present invention describes a method for determining whether a
CC   compound induces a toxic effect on a tissue or cell. The method comprises
CC   preparing a gene expression profile of a tissue or cell sample exposed to
CC   the compound, and comparing the gene expression profile to a database
CC   comprising data or information on the Tox mean and non-Tox mean value.
CC   The method is useful for predicting or identifying at least one toxic
CC   effect, particularly hepatotoxicity, of a test or unknown compound. The
CC   genes listed in the specification are useful as diagnostic or toxicity
CC   markers for the prediction or identification of the physiological state
CC   of tissue or cell sample that has been exposed to a compound, or to
CC   identify or predict the toxic effects of a compound or an agent. These
CC   may also be used as markers for monitoring toxicity progression or for
CC   drug screening. The present sequence represents a primary rat hepatocyte
CC   toxicity modelling related gene sequence from the present invention.
XX
SQ   Sequence 8739 BP; 2288 A; 2115 C; 2109 G; 2227 T; 0 U; 0 Other;

  Query Match             80.9%;  Score 2909;  DB 9;  Length 8739;
  Best Local Similarity   88.9%;  Pred. No. 0;
  Matches 3168;  Conservative    0;  Mismatches  395;  Indels    2;  Gaps    2;

Qy         15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74
              || |||    |  ||   | ||  |  |     | | |  ||||  ||    | || ||
```





```
Qy      2234 ACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGC 2293
             ||||||||| |||||||||| |||||  |||||||||||||| || ||||||||||||||||
Db      2525 ACAATGCCTTTTTGGCTTCGCTTACGTCCCAGAAGAAGGTGGAGTTTGTCCTTAAGCAGC 2584

Qy      2294 TGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCC 2353
             | ||| ||||||| |||||||||||||| ||||||||||||  | || |||||||| || |
Db      2585 TTCGATTAATGCAATCATCTCAGAGCACGTCCAAGCTCACTCTGACTCCATGGGTTGGTC 2644

Qy      2354 TTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTT 2413
             |||||||||||||||| || |||||||||||||| |||||||| ||||| ||||||||||
Db      2645 TTCGGAAGATCAATGTATCTTACTGGTGCTGGGAGGATATGTCTCCATTCACAAATAGTT 2704

Qy      2414 TACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAAC 2473
             | || |||||||||||| |||||||||||||||| || |||||||| || |||| | |
Db      2705 TGCTGCAGTGGATGCCATCTGAGCCCAGTGATGCCGGCTTCTGTGGGATCTTGTCAGAGC 2764

Qy      2474 CCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTG 2533
             | ||||||||||||||||||||||||||||||||||||||||||||| ||||| ||||
Db      2765 CTAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGAAGTGTTTGTG 2824

Qy      2534 AAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCAT 2593
             |||||||||||||||||||| |||||||||||||||| |||||| || |||||||||
Db      2825 AAAGGCCTGCAAACCACAGTGCCAAGCAGTGCCGGACACCCTGTGCCCTGCGGACAGCAT 2884

Qy      2594 GTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTG 2653
             |||| || |||||||| ||||||| || ||||||||||||||||||||||||||||||| |
Db      2885 GTGGTGAGTGCACCAGCAGCAGCTCCGAATGCATGTGGTGCAGCAACATGAAGCAGTGCG 2944

Qy      2654 TGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGA 2713
             | |||||||||||||| || |||||||| ||||| |||||||||||||| |||||||| |
Db      2945 TAGACTCCAATGCCTACGTAGCCTCCTTTCCTTTCGGCCAGTGTATGGAGTGGTATACAA 3004

Qy      2714 TGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGC 2773
             |||||| |||||| |||||||||| || |||||||||||||| || ||||||||||||||
Db      3005 TGAGCAGCTGCCCACCTGAAAATTGCTCTGGCTACTGTACCTGCAGCCATTGCTTGGAGC 3064

Qy      2774 AACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGG 2833
             | |||||||||||| |||||||||||||| |||||||||| ||||| ||||| || ||||
Db      3065 AGCCAGGCTGTGGTTGGTGTACTGATCCTAGCAATACTGGGAAAGGCAAATGTATTGAGG 3124

Qy      2834 GTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCAC 2893
             |   |||||||||||| |||||||||||||| || ||| |||||||||| | || ||||
Db      3125 GAAGCTATAAAGGACCCGTGAAGATGCCTTCACATGCCTCTACAGGAAATGTGTACCCAC 3184

Qy      2894 AGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTC 2953
             ||||||| || || |||||||||||||||||||||||||||||||||||||||||| ||||
Db      3185 AGCCCCTTCTGAACTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTTATTC 3244

Qy      2954 ACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGA 3013
             |||||||||||||||| ||||| || ||||| || |||||||| |||||||||||||||||
Db      3245 ACTGTCCAGCTTGCCAGTGCAATGGACACAGCAAGTGCATCAACCAGAGCATCTGTGAGA 3304

Qy      3014 AGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTG 3073
             |||||||| |||||||||||||||||||||||||||||||||||||||||||||| ||||
Db      3305 AGTGTGAGGACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTATGGTG 3364

Qy      3074 ATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCA 3133
             | || || |||||||| |||||||||||||||||||||||||||||||| |||||||| ||||||||
Db      3365 ACCCGACTAATGGAGGCAAATGTCAGCCATGCAAGTGCAACGGGCACGCATCACTGTGCA 3424

Qy      3134 ACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGC 3193
             |||||||||| |||||||||||||| |||||||||||||||||||||| || ||||||||||
Db      3425 ACACCAACACTGGCAAGTGCTTCTGTACCACCAAGGGCGTCAAGGGAGAGGAGTGCCAGC 3484

Qy      3194 TATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTC 3253
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db      3485 TATGTGAGGTAGAAAATCGATACCAGGGAAACCCTCTCAAAGGAACATGTTACTATACCC 3544

Qy      3254 TTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAG 3313
             |||| |||||||||||||||| ||||| || |||||||||| || ||||||||||||||
Db      3545 TTCTCATTGACTATCAGTTCACTTTTAGCCTGTCCCAGGAAGACGACCGCTATTACACAG 3604

Qy      3314 CTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATG 3373
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3605 CCATCAATTTTGTGGCTACTCCTGATGAACAAAACAGGGATTTGGACATGTTCATCAATG 3664

Qy      3374 CCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGG 3433
             ||||||| || |||||||||||||||||||||||||||| |||||||| ||||||||||||
Db      3665 CCTCCAAAAACTTCAACCTCAACATCACCTGGGCCACCAGTTTCCCAGCTGGAACCCAGA 3724

Qy      3434 CTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCT 3493
             |||||||||| |||||||||||||||||||||||||||||| ||||| |||||||||| ||||
Db      3725 CTGGAGAAGAGGTGCCTGTTGTTTCAAAAACCAACATCAAGGAATACAAAGATAGCTTCT 3784

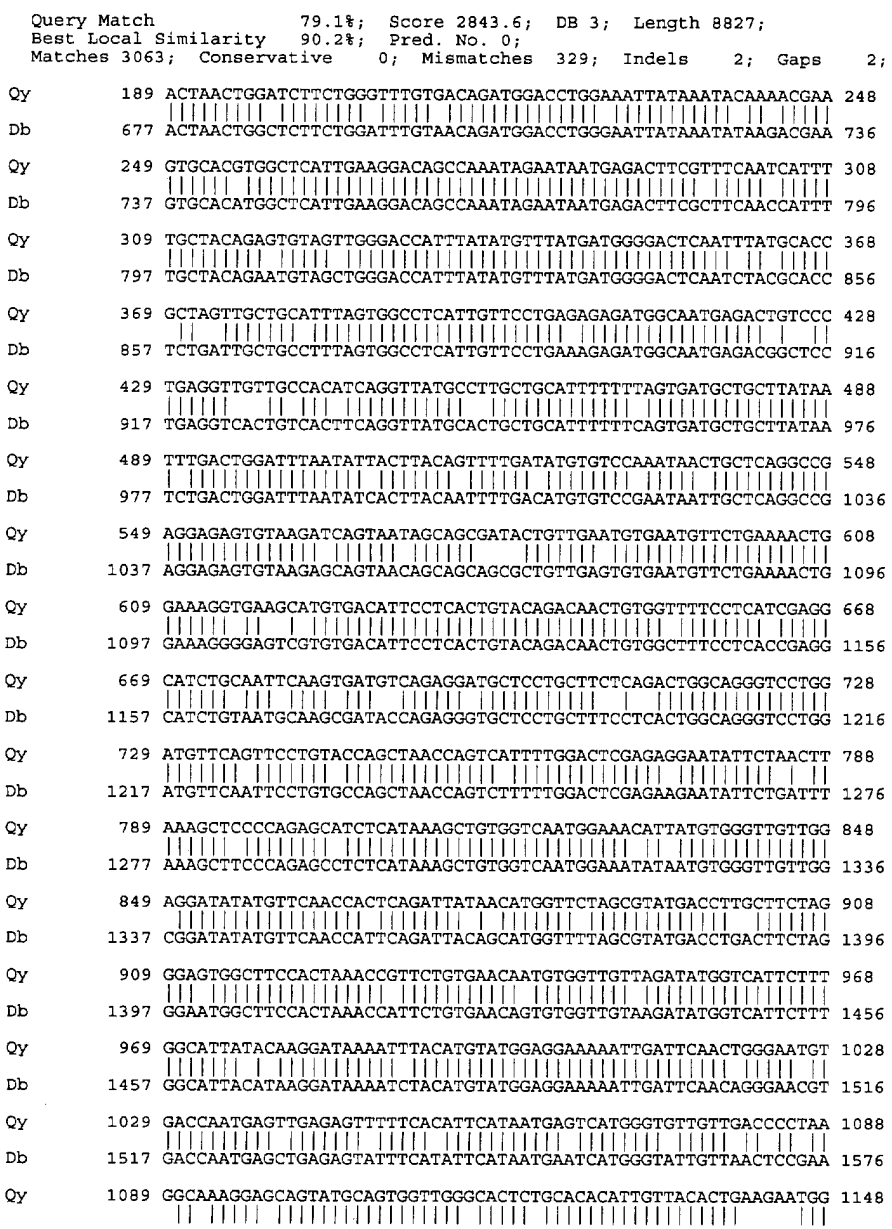
Qy      3494 CTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTA 3553
             |||||||||| |||||||||||||||| ||||| || ||||||||||||||||||||||
Db      3785 CTAATGAGAAATTTGATTTTCGCAACCATCCAAACATTACTTTCTTTGTTTATGTCAGTA 3844

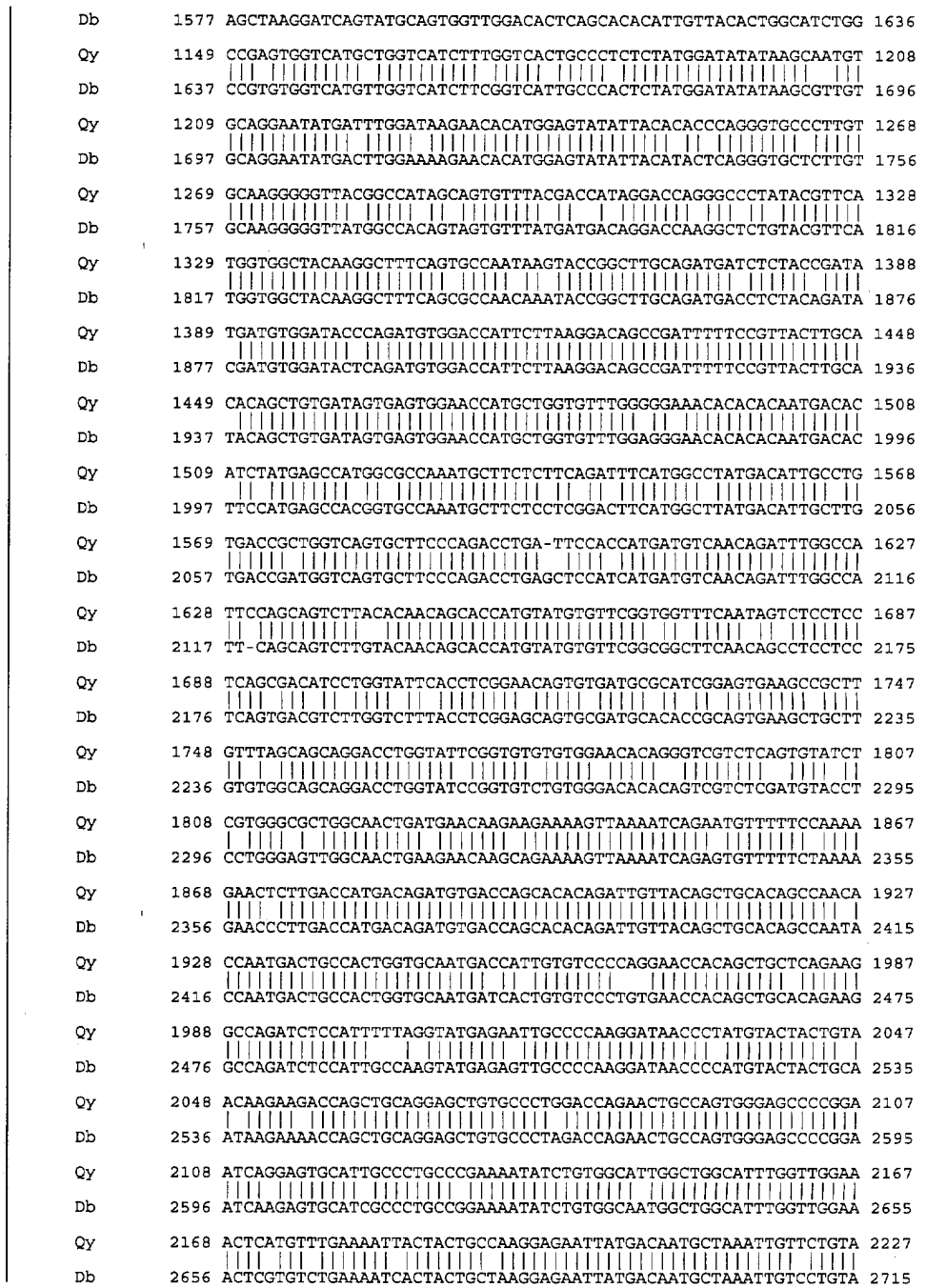
Qy      3554 ATTTCACCTGGCCCATCAAAATTCA 3578
             ||||||| ||||||||||||||||
Db      3845 ATTTCACTTGGCCCATCAAAATTCA 3869
```

RESULT 11
AAZ91916
ID AAZ91916 standard; cDNA; 8827 BP.
XX
AC AAZ91916;
XX
DT 08-JUN-2000 (first entry)
XX
DE Wild type (C57BL/6J) murine mahogany protein coding sequence.
XX
KW Mahogany gene; mouse; mg gene; regulatory defect; gene therapy; obesity;
KW weight regulation; cell therapy; body weight disorder; cachexia;
KW anorexia; hyperpigmentation; increased metabolic rate disorder;
KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.
XX
OS Mus sp.
XX
PN WO200005373-A2.
XX
PD 03-FEB-2000.
XX
PF 21-JUL-1999; 99WO-US016484.
XX
PR 21-JUL-1998; 98US-0093630P.
PR 20-OCT-1998; 98US-0104978P.
PR 05-FEB-1999; 99US-00245041.
XX
PA (MILL-) MILLENIUM PHARM INC.
XX
PI Moore K, Nagle DL;
XX
DR WPI; 2000-195103/17.
DR P-PSDB; AAY81803.
XX
PT New human and murine mahogany genes, useful, e.g. for diagnosis and
PT treatment of body weight disorders.
XX
PS Claim 1; Fig 2a; 188pp; English.
XX
CC This sequence represents a murine mahogany gene of the invention. The
CC mahogany genes are used: (i) to produce recombinant mahogany (mg)
CC proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming
CC therapeutics; (iii) as a source of diagnostic probes and primers for
CC detecting expression of mg genes or mutations, regulatory defects, in
CC this gene, or for isolation of related sequences; and (iv) in (cell-
CC based) gene therapy. (II) are used to raise specific antibodies (Ab); to
CC identify other (extra)cellular products involved in weight regulation,
CC and to screen for agents that disrupt interaction between (II) and other
CC macromolecules. The Ab are used to detect abnormal levels (or function)
CC of (II) (for diagnosis, prognosis or monitoring of treatment); to
CC evaluate (II)-expressing cells intended for cell therapy, and as
CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the
CC mg polypeptide) are used to identify agents (A) that modulate mg

CC activity. (A) are potentially useful for the treatment of body weight
CC disorders, particularly obesity, cachexia or anorexia, or other
CC conditions associated with the mg gene such as hyperpigmentation,
CC hyperphagia and disorders that result in increased metabolic rate
XX
SQ Sequence 8827 BP; 2304 A; 2077 C; 2162 G; 2284 T; 0 U; 0 Other;

Query Match 79.1%; Score 2843.6; DB 3; Length 8827;
Best Local Similarity 90.2%; Pred. No. 0;
Matches 3063; Conservative 0; Mismatches 329; Indels 2; Gaps 2;





```
Qy       2228 GGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTA 2287
Db       2716 GGAACCACAATGCCTTTTTGGCTTCCCTCACATCCCAGAAGAAGGTGGAGTTTGTCCTTA 2775

Qy       2288 AGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGG 2347
Db       2776 AGCAGCTTCGATTAATGCAATCATCTCAAAGTATGTCCAAGCTCACTCTGACTCCATGGG 2835

Qy       2348 TCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAA 2407
Db       2836 TTGGTCTTCGGAAGATCAATGTGTCTTACTGGTGCTGGGAGGATATGTCTCCATTCACAA 2895

Qy       2408 ATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAAATTTTAT 2467
Db       2896 ATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGTGATGCTGGCTTCTGTGGGATCTTGT 2955

Qy       2468 CAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTG 2527
Db       2956 CAGAGCCTAGTACTCGGGGATTAAAGGCTGCAACCTGCATCAACCCTCTCAATGGCAGCG 3015

Qy       2528 TCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGA 2587
Db       3016 TCTGTGAAAGGCCTGCAAACCACAGTGCCAAGCAGTGCCGGACACCATGTGCCCTGCGGA 3075

Qy       2588 CAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGC 2647
Db       3076 CAGCGTGTGGCGAGTGCACTAGCAGCAGCTCGGAGTGCATGTGGTGCAGTAACATGAAGC 3135

Qy       2648 AGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGT 2707
Db       3136 AGTGTGTGGACTCCAATGCCTACGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGT 3195

Qy       2708 ATACGATGAGCACCTGCCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCT 2767
Db       3196 ATACGATGAGCAGCTGCCCACCTGAAAATTGCTCTGGCTACTGTACCTGCAGCCATTGCT 3255

Qy       2768 TGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCA 2827
Db       3256 TGGAGCAGCCAGGCTGTGGTTGGTGTACTGATCCTAGCAATACTGGGAAAGGAAAATGTA 3315

Qy       2828 TAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCT 2887
Db       3316 TTGAGGGCAGCTATAAAGGACCTGTGAAGATGCCGTCACAGGCCTCTGCAGGAAATGTGT 3375

Qy       2888 ATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTT 2947
Db       3376 ATCCACAGCCCCTTCTGAACTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTT 3435

Qy       2948 TCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCT 3007
Db       3436 TCATTCACTGTCCAGCTTGCCAGTGCAACGGACACAGCAAATGCATCAACCAGAGTATCT 3495

Qy       3008 GTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCT 3067
Db       3496 GTGAGAAGTGTGAGGACCTGACCACGGGCAAGCACTGCGAGACCTGCATATCTGGCTTCT 3555

Qy       3068 ACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTC 3127
Db       3556 ATGGTGACCCGACTAATGGAGGCAAATGTCAGCCATGCAAGTGCAATGGGCACGCATCAC 3615

Qy       3128 TGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGT 3187
Db       3616 TGTGCAACACCAACACCGGCAAGTGCTTCTGTACCACCAAAGGTGTCAAGGGGGACGAGT 3675

Qy       3188 GCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATT 3247
Db       3676 GCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAAAGGAACATGCTACT 3735

Qy       3248 ATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATT 3307
Db       3736 ATACCCTTCTCATTGACTATCAGTTCACCTTTAGCCTGTCCCAGGAAGACGACCGCTACT 3795

Qy       3308 ACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCA 3367
Db       3796 ACACAGCCATCAACTTTGTGGCTACTCCTGATGAACAAAACAGGGATTTGGACATGTTCA 3855

Qy       3368 TCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAA 3427
Db       3856 TCAATGCCTCCAAAAACTTCAACCTCAACATCACCTGGGCCACCAGCTTCCCAGCCGGAA 3915

Qy       3428 CCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATA 3487
Db       3916 CCCAGACTGGAGAAGAGGTGCCTGTTGTTTCAAAAACCAACATCAAGGAATACAAAGATA 3975

Qy       3488 GTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATG 3547
Db       3976 GCTTCTCTAATGAGAAATTTGATTTTCGCAACGATCCAAACATCACTTTCTTTGTTTATG 4035

Qy       3548 TCAGTAATTTCACCTGGCCCATCAAAATTCAGGT 3581
Db       4036 TCAGTAATTTCACTTGGCCCATCAAAATTCAGAT 4069
```

RESULT 12
AAS72660
ID AAS72660 standard; cDNA; 3490 BP.
XX
AC AAS72660;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #8464.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG08473.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 8464; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations

```
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 3490 BP; 931 A; 772 C; 858 G; 929 T; 0 U; 0 Other;

  Query Match             77.6%;  Score 2790.4;  DB 5;  Length 3490;
  Best Local Similarity   94.2%;  Pred. No. 0;
  Matches 3238;  Conservative    0;  Mismatches   51;  Indels  147;  Gaps   27;

Qy        187 AGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACG 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 AGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACG 81

Qy        247 AAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCAT 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 AAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCAT 141

Qy        307 TTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCA 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 TTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCA 201

Qy        367 CCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTC 426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 CCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTC 261

Qy        427 CCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTTAGTGATGCTGCTTAT 486
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 CCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTTAGTGATGCTGCTTAT 321

Qy        487 AATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGC 546
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 AATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGC 381

Qy        547 CGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAAC 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        382 CGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAAC 441

Qy        607 TGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGA 666
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        442 TGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGA 501

Qy        667 GGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCT 726
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        502 GGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCT 561

Qy        727 GGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTT-GGACTCGAGAGGAATATTCTAA 785
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db        562 GGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGGACTCGAGAGGAATATTCTAA 621

Qy        786 CTT--AAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTT 843
              |||  ||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        622 CTTTAAAGCTCCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTT 681

Qy        844 GTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCT 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        682 GTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCT 741

Qy        904 TCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCAT 963
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        742 TCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCAT 801

Qy        964 TCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGG 1023
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        802 TCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGG 861

Qy       1024 AATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACC 1083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        862 AATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACC 921

Qy       1084 CCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAG 1143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        922 CCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAG 981

Qy       1144 AATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGC 1203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        982 AATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGC 1041

Qy       1204 AATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCC 1263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1042 AATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCC 1101

Qy       1264 CTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATAC 1323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1102 CTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATAC 1161

Qy       1324 GTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTAC 1383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1162 GTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTAC 1221

Qy       1384 CGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTAC 1443
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1222 CGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTAC 1281

Qy       1444 TTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAAT 1503
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1282 TTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGAGGAAACACACACAAT 1341

Qy       1504 GACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATT 1563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1342 GACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATT 1401

Qy       1564 GCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGAT-TCCACCATGATGTCAACAGATTT 1622
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db       1402 GCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTT 1461

Qy       1623 GGCCATTCCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCT 1682
              ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1462 GGCCATT-CAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCT 1520

Qy       1683 CCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGC 1742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1521 CCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGC 1580

Qy       1743 CGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTG 1802
              ||||||||||||||||||||||||||||||
Db       1581 CGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGT-------------------------- 1614

Qy       1803 TATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTC 1862

Db       1615 ------------------------------------------------------------ 1614

Qy       1863 CAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGC 1922
                      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1615 --------CTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGC 1666

Qy       1923 CAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTC 1982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1667 CAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTC 1726

Qy       1983 AGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTA 2042
              |||||||          ||||||||||||||||||||||||||||| ||||||||
Db       1727 AGAAGGCC----------AGGTATGAGAATTGCCCCAAGGATAACCCCATGTACTA 1772

Qy       2043 CTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCC 2102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1773 CTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCC 1832

Qy       2103 CCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGT 2162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1833 CCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGT 1892
```

```
Qy       2163 TGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTT 2222
Db       1893 TGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTT 1952

Qy       2223 CTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTA-GAATTTG 2281
Db       1953 CTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGGAATTTG 2012

Qy       2282 TCCTTAAG--CAGCTGCGAATAATGCA-GTCATCTCAGA-GCATGTCCAA-GCTCACCTT 2336
Db       2013 TCCTTAAGGCAGGCTGCGAATAATGCAGGTCATCTCAGAGGCATGTCCAAGGCTCACCTT 2072

Qy       2337 AACCCCATGGG---TCGGCCTTCGGAAGATCAA-TGTGTCCTACTGGTGCTGGGAAGATA 2392
Db       2073 AACCCCATGGGGTCGGGCCTTCGGGAAGATCAATTGTGTCCTACTGGTGCTGGGAAGATA 2132

Qy       2393 TGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGAT 2452
Db       2133 TGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGAT 2192

Qy       2453 TCTGTGGAATTTT-----ATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCAT 2507
Db       2193 TCTGTGGGAATTTTATTCAGGAACCCAGTTACTTCGGGGACTGAAGGCTGCAACCTGCAT 2252

Qy       2508 --CAACCCACTCAATGGTAGTGTCTGTGAAA--GGCCTGCAAACCACAGT-GCTAAGCAG 2562
Db       2253 TCAACCCACTTCAATGGTAGTGTCTGTGAAAAGGCCTGGCAAACCACAGTGGTTAAGCAG 2312

Qy       2563 TGCCGGACACCATGTGCC----TTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTC 2618
Db       2313 TGCCGGACACCATGGTGCCCTTTGAGGGACAGCATGTGGAGATTGCACCAGCGGCAGCTC 2372

Qy       2619 TGAGT-GCATGTGGTGCAGCAACATGAA-GCAGTGTGTGGACTCCAATGCCTATGTGGCC 2676
Db       2373 TGAGTGGCATGTGGTGCAGCAACATGAAGGCAGTGTGTGGACTCCAATGCCTATGTGGCC 2432

Qy       2677 TCCTTCCCTTTT-GGCCAGTGTATGGAATGGTATACGATGA-GCACCTGCCCCCCTGAAA 2734
Db       2433 TCCTTCCCTTTTGGGCCAGTGTATGGAATGGTATACGATGAGGCACCTGCCCCCCTGAAA 2492

Qy       2735 ATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCA-GGCTGTGGCTGGTGT 2793
Db       2493 ATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGGCTGTGGCTGGTGT 2552

Qy       2794 ACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTG 2853
Db       2553 ACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTG 2612

Qy       2854 AAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGC 2913
Db       2613 AAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGC 2672

Qy       2914 ATGTGTCTAGAGGACAGCAGATACAACTGG--TCTTTCATTCACTGTCCAGCTTGCCAAT 2971
Db       2673 ATGTGTCTAGAGGACAGCAGATACAACTGGGTCTTTCATTTCACTGTCCAGTTTGCCAAT 2732

Qy       2972 GCAAC-GGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACC 3030
Db       2733 GGAACGGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACC 2792

Qy       3031 ACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTAC-GGTGATCCCACCAA-TGGAG 3088
Db       2793 ACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGGTGATCCCACCAATTGGAG 2852

Qy       3089 GGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCA 3148
Db       2853 GGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCA 2912

Qy       3149 AGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAA 3208
Db       2913 AGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAA 2972

Qy       3209 ATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATC 3268
Db       2973 ATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATC 3032

Qy       3269 AGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGG 3328
Db       3033 AGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGG 3092

Qy       3329 CTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCA 3388
Db       3093 CTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCA 3152

Qy       3389 ACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGC 3448
Db       3153 ACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGC 3212

Qy       3449 CTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTG 3508
Db       3213 CTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTG 3272

Qy       3509 ATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCA 3568
Db       3273 ATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCA 3332

Qy       3569 TCAAAATTCAGGTGCA 3584
Db       3333 TCAAAATTCAGGTGAA 3348
```

RESULT 13
AAZ91922
ID AAZ91922 standard; cDNA; 2625 BP.
XX
AC AAZ91922;
XX
DT 08-JUN-2000 (first entry)
XX
DE Human mahogany protein coding sequence #4.
XX
KW Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;
KW weight regulation; cell therapy; body weight disorder; cachexia;
KW anorexia; hyperpigmentation; increased metabolic rate disorder;
KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.
XX
OS Homo sapiens.
XX
PN WO200005373-A2.
XX
PD 03-FEB-2000.
XX
PF 21-JUL-1999; 99WO-US016484.
XX
PR 21-JUL-1998; 98US-0093630P.
PR 20-OCT-1998; 98US-0104978P.
PR 05-FEB-1999; 99US-00245041.
XX
PA (MILL-) MILLENIUM PHARM INC.
XX
PI Moore K, Nagle DL;
XX
DR WPI; 2000-195103/17.
DR P-PSDB; AAY81809.
XX
PT New human and murine mahogany genes, useful, e.g. for diagnosis and
PT treatment of body weight disorders.
XX
PS Claim 1; Fig 20a; 188pp; English.
XX
CC This sequence represents a human mahogany gene of the invention. The
CC mahogany genes are used: (i) to produce recombinant mahogany (mg)
CC proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming
CC therapeutics; (iii) as a source of diagnostic probes and primers for
CC detecting expression of mg genes or mutations, regulatory defects, in
CC this gene, or for isolation of related sequences; and (iv) in (cell-

```
CC   based) gene therapy. (II) are used to raise specific antibodies (Ab); to
CC   identify other (extra)cellular products involved in weight regulation,
CC   and to screen for agents that disrupt interaction between (II) and other
CC   macromolecules. The Ab are used to detect abnormal levels (or function)
CC   of (II) (for diagnosis, prognosis or monitoring of treatment); to
CC   evaluate (II)-expressing cells intended for cell therapy, and as
CC   therapeutic mg inhibitors. Cells that express the mg gene (or contain the
CC   mg polypeptide) are used to identify agents (A) that modulate mg
CC   activity. (A) are potentially useful for the treatment of body weight
CC   disorders, particularly obesity, cachexia or anorexia, or other
CC   conditions associated with the mg gene such as hyperpigmentation,
CC   hyperphagia and disorders that result in increased metabolic rate
XX
SQ   Sequence 2625 BP; 660 A; 605 C; 694 G; 666 T; 0 U; 0 Other;

  Query Match             56.4%;  Score 2028.8;  DB 3;  Length 2625;
  Best Local Similarity   97.0%;  Pred. No. 0;
  Matches 2088;  Conservative    0;  Mismatches   62;  Indels    2;  Gaps    2;

Qy         15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74
Db        234 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAA 293

Qy         75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134
Db        294 GGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 353

Qy        135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194
Db        354 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 413

Qy        195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254
Db        414 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 473

Qy        255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314
Db        474 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 533

Qy        315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374
Db        534 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 593

Qy        375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434
Db        594 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 653

Qy        435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494
Db        654 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 713

Qy        495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554
Db        714 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 773

Qy        555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614
Db        774 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 833

Qy        615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 674
Db        834 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 893

Qy        675 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 734
Db        894 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 953

Qy        735 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 794
Db        954 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 1013

Qy        795 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 854
Db       1014 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 1073

Qy        855 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 914
Db       1074 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 1133

Qy        915 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 974
Db       1134 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 1193

Qy        975 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1034
Db       1194 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1253

Qy       1035 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1094
Db       1254 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1313

Qy       1095 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1154
Db       1314 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1373

Qy       1155 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1214
Db       1374 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1433

Qy       1215 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1274
Db       1434 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1493

Qy       1275 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1334
Db       1494 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1553

Qy       1335 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1394
Db       1554 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1613

Qy       1395 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1454
Db       1614 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1673

Qy       1455 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1514
Db       1674 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1733

Qy       1515 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1574
Db       1734 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1793

Qy       1575 CTGGTCAGTGCTTCCCAGACCTGAT-TCCACCATGATGTCAACAGATTTGGCCATTCCAG 1633
Db       1794 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATT-CAG 1852

Qy       1634 CAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCG 1693
Db       1853 CAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCG 1912

Qy       1694 ACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAG 1753
Db       1913 ACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAG 1972

Qy       1754 CAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGG 1813
Db       1973 CAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGG 2032

Qy       1814 CGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTC 1873
Db       2033 CGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTC 2092

Qy       1874 TTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATG 1933
Db       2093 TTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATG 2152
```

```
Qy       1934 ACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGA 1993
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2153 ACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGA 2212

Qy       1994 TCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGA 2053
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2213 TCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGA 2272

Qy       2054 AGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGG 2113
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2273 AGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGG 2332

Qy       2114 AGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGG 2165
             ||||||||||||||||||  |    ||  |  |  ||  |   |  ||  ||
Db       2333 AGTGCATTGCCCTGCCCGGTAGGCCTTGCAGGGTCATCTTGGTGTGTGTGGG 2384
```

RESULT 14
AAZ91917
ID AAZ91917 standard; cDNA; 2419 BP.
XX
AC AAZ91917;
XX
DT 08-JUN-2000 (first entry)
XX
DE Murine mahogany protein coding sequence akm1003.
XX
KW Mahogany gene; mouse; mg gene; regulatory defect; gene therapy; obesity;
KW weight regulation; cell therapy; body weight disorder; cachexia;
KW anorexia; hyperpigmentation; increased metabolic rate disorder;
KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.
XX
OS Mus sp.
XX
PN WO200005373-A2.
XX
PD 03-FEB-2000.
XX
PF 21-JUL-1999; 99WO-US016484.
XX
PR 21-JUL-1998; 98US-0093630P.
PR 20-OCT-1998; 98US-0104978P.
PR 05-FEB-1999; 99US-00245041.
XX
PA (MILL-) MILLENIUM PHARM INC.
XX
PI Moore K, Nagle DL;
XX
DR WPI; 2000-195103/17.
DR P-PSDB; AAY81804.
XX
PT New human and murine mahogany genes, useful, e.g. for diagnosis and
PT treatment of body weight disorders.
XX
PS Claim 1; Fig 8a; 188pp; English.
XX
CC This sequence represents a murine mahogany gene of the invention. The
CC mahogany genes are used: (i) to produce recombinant mahogany (mg)
CC proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming
CC therapeutics; (iii) as a source of diagnostic probes and primers for
CC detecting expression of mg genes or mutations, regulatory defects, in
CC this gene, or for isolation of related sequences; and (iv) in (cell-
CC based) gene therapy. (II) are used to raise specific antibodies (Ab); to
CC identify other (extra)cellular products involved in weight regulation,
CC and to screen for agents that disrupt interaction between (II) and other
CC macromolecules. The Ab are used to detect abnormal levels (or function)
CC of (II) (for diagnosis, prognosis or monitoring of treatment); to
CC evaluate (II)-expressing cells intended for cell therapy, and as
CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the
CC mg polypeptide) are used to identify agents (A) that modulate mg
CC activity. (A) are potentially useful for the treatment of body weight
CC disorders, particularly obesity, cachexia or anorexia, or other
CC conditions associated with the mg gene such as hyperpigmentation,
CC hyperphagia and disorders that result in increased metabolic rate
XX
SQ Sequence 2419 BP; 578 A; 567 C; 680 G; 594 T; 0 U; 0 Other;

Query Match 39.7%; Score 1428.8; DB 3; Length 2419;
Best Local Similarity 89.4%; Pred. No. 0;
Matches 1560; Conservative 0; Mismatches 182; Indels 2; Gaps 2;

```
Qy        189 ACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAA 248
              |||||||| |||||||| ||||| |||||||||||||| |||||||||| || |||||
Db        677 ACTAACTGGCTCTTCTGGATTTGTAACAGATGGACCTGGGAATTATAAATATAAGACGAA 736

Qy        249 GTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTT 308
              |||||| |||||||||||||||||||||||||||||||||||||||||| ||||| |||||
Db        737 GTGCACATGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGCTTCAACCATTT 796

Qy        309 TGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACC 368
              |||||||| ||||| |||||||||||||||||||||||||||||||||| || |||||
Db        797 TGCTACAGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATCTACGCACC 856

Qy        369 GCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCC 428
              ||  |||||||  |||||||||||||||||||||| ||||||||||||||||| |  ||
Db        857 TCTGATTGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCTCC 916

Qy        429 TGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAA 488
              ||||||   || ||| |||||||||||  |||||||||||| |||||||||||||||||
Db        917 TGAGGTCACTGTCACTTCAGGTTATGCACTGCTGCATTTTTTCAGTGATGCTGCTTATAA 976

Qy        489 TTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCG 548
              | |||||||||||||||| ||||||| |||||| |||||||| ||||| ||||||||||
Db        977 TCTGACTGGATTTAATATCACTTACAATTTTGACATGTGTCCGAATAATTGCTCAGGCCG 1036

Qy        549 AGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTG 608
              |||||||||||||| |||||| ||||||    ||||||| |||||||||||||||||||
Db       1037 AGGAGAGTGTAAGAGCAGTAACAGCAGCAGCGCTGTTGAGTGTGAATGTTCTGAAAACTG 1096

Qy        609 GAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGG 668
              |||||  ||   | |||||||||||||||||||||||||||||||| |||||||| |||||
Db       1097 GAAAGGGGAGTCGTGTGACATTCCTCACTGTACAGACAACTGTGGCTTTCCTCACCGAGG 1156

Qy        669 CATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGG 728
              |||||| ||| |||| ||| |||||| ||||||||||  |  ||||||||||||||||
Db       1157 CATCTGTAATGCAAGCGATACCAGAGGGTGCTCCTGCTTTCCTCACTGGCAGGGTCCTGG 1216

Qy        729 ATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTT 788
              ||||||| ||||||| |||||||||||||| |||||||||||||| |||||||| | ||
Db       1217 ATGTTCAATTCCTGTGCCAGCTAACCAGTCTTTTTGGACTCGAGAAGAATATTCTGATTT 1276

Qy        789 AAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGG 848
              |||||| |||||||| |||||||||||||||||||||||||| || ||||||||||||||
Db       1277 AAAGCTTCCCAGAGCCTCTCATAAAGCTGTGGTCAATGGAAATATAATGTGGGTTGTTGG 1336

Qy        849 AGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAG 908
              |||||||||||||||| |||||||| ||||||| |||||||||||| |||||||||||||
Db       1337 CGGATATATGTTCAACCATTCAGATTACAGCATGGTTTTAGCGTATGACCTGACTTCTAG 1396

Qy        909 GGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTT 968
              ||| |||||||||||||| |||||||||| |||||||| ||||||||||||||||||||
Db       1397 GGAATGGCTTCCACTAAACCATTCTGTGAACAGTGTGGTTGTAAGATATGGTCATTCTTT 1456

Qy        969 GGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGT 1028
              |||||| | |||||||||| |||||||||||||||||||||||||||||||||| |||| ||
Db       1457 GGCATTACATAAGGATAAAATCTACATGTATGGAGGAAAAATTGATTCAACAGGGAACGT 1516

Qy       1029 GACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAA 1088
              ||||||||| ||||||| ||||| |||||||||| |||||||| ||||| || || ||
Db       1517 GACCAATGAGCTGAGAGTATTTCATATTCATAATGAATCATGGGTATTGTTAACTCCGAA 1576

Qy       1089 GGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGG 1148
              || ||||| |||||||||||||||| |||| |||||||||||||||||  |||
Db       1577 AGCTAAGGATCAGTATGCAGTGGTTGGACACTCAGCACACATTGTTACACTGGCATCTGG 1636
```

```
Qy       1149 CCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGT 1208
              ||| |||||||||| |||||||||| ||||| ||||| ||||||||||||||||   |||
Db       1637 CCGTGTGGTCATGTTGGTCATCTTCGGTCATTGCCCACTCTATGGATATATAAGCGTTGT 1696

Qy       1209 GCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGT 1268
              |||||||||||| ||||| |||||||||||||||||||||||| || |||||||| |||||
Db       1697 GCAGGAATATGACTTGGAAAAGAACACATGGAGTATATTACATACTCAGGGTGCTCTTGT 1756

Qy       1269 GCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCA 1328
              |||||||||||| ||||| || |||||||| ||   | ||||||| ||| || ||||||||
Db       1757 GCAAGGGGGTTATGGCCACAGTAGTGTTTATGATGACAGGACCAAGGCTCTGTACGTTCA 1816

Qy       1329 TGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATA 1388
              |||||||||||||||||||| ||||| || |||||||||||||||| |||||| ||||
Db       1817 TGGTGGCTACAAGGCTTTCAGCGCCAACAAATACCGGCTTGCAGATGACCTCTACAGATA 1876

Qy       1389 TGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCA 1448
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       1877 CGATGTGGATACTCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCA 1936

Qy       1449 CACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACAC 1508
              |||||||||||||||||||||||||||||||||||| || ||||||||||||||||||
Db       1937 TACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGAGGGAACACACACAATGACAC 1996

Qy       1509 ATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTG 1568
              || |||||||| || |||||||||||||| || || |||||||| |||||||||| ||
Db       1997 TTCCATGAGCCACGGTGCCAAATGCTTCTCCTCGGACTTCATGGCTTATGACATTGCTTG 2056

Qy       1569 TGACCGCTGGTCAGTGCTTCCCAGACCTGA-TTCCACCATGATGTCAACAGATTTGGCCA 1627
              |||||| |||||||||||||||||||||   |||| ||||||||||||||||||||||
Db       2057 TGACCGATGGTCAGTGCTTCCCAGACCTGAGCTCCATCATGATGTCAACAGATTTGGCCA 2116

Qy       1628 TTCCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCC 1687
              || |||||||||| |||||||||||||||||||||||| || ||||| || |||||||
Db       2117 TT-CAGCAGTCTTGTACAACAGCACCATGTATGTGTTCGGCGGCTTCAACAGCCTCCTCC 2175

Qy       1688 TCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTT 1747
              |||| ||| || |||| || |||||||| ||||| ||||| || || |||||||| ||||
Db       2176 TCAGTGACGTCTTGGTCTTTACCTCGGAGCAGTGCGATGCACACCGCAGTGAAGCTGCTT 2235

Qy       1748 GTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCT 1807
              || | |||||||||||||||| |||||| ||||| |||||   |||||||  |||| ||
Db       2236 GTGTGGCAGCAGGACCTGGTATCCGGTGTCTGTGGGACACACAGTCGTCTCGATGTACCT 2295

Qy       1808 CGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAA 1867
              | |||| | |||||||||| ||||||| |||||||||||||||||| |||||||| ||||
Db       2296 CCTGGGAGTTGGCAACTGAAGAACAAGCAGAAAAGTTAAAATCAGAGTGTTTTTCTAAAA 2355

Qy       1868 GAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACA 1927
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       2356 GAACCCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAATA 2415

Qy       1928 CCAA 1931
              ||||
Db       2416 CCAA 2419
```

RESULT 15
AAZ91919
ID AAZ91919 standard; cDNA; 6373 BP.
XX
AC AAZ91919;
XX
DT 08-JUN-2000 (first entry)
XX
DE Human mahogany protein coding sequence #1.
XX
KW Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;
KW weight regulation; cell therapy; body weight disorder; cachexia;
KW anorexia; hyperpigmentation; increased metabolic rate disorder;
KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT misc_difference 448
FT /*tag= a
FT /note= "represented in specification as :"
FT misc_difference 615
FT /*tag= b
FT /note= "represented in specification as :"
FT misc_difference 683
FT /*tag= c
FT /note= "represented in specification as :"
XX
PN WO200005373-A2.
XX
PD 03-FEB-2000.
XX
PF 21-JUL-1999; 99WO-US016484.
XX
PR 21-JUL-1998; 98US-0093630P.
PR 20-OCT-1998; 98US-0104978P.
PR 05-FEB-1999; 99US-00245041.
XX
PA (MILL-) MILLENIUM PHARM INC.
XX
PI Moore K, Nagle DL;
XX
DR WPI; 2000-195103/17.
DR P-PSDB; AAY81806.
XX
PT New human and murine mahogany genes, useful, e.g. for diagnosis and
PT treatment of body weight disorders.
XX
PS Claim 1; Fig 10a; 188pp; English.
XX
CC This sequence represents a human mahogany gene of the invention. The
CC mahogany genes are used: (i) to produce recombinant mahogany (mg)
CC proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming
CC therapeutics; (iii) as a source of diagnostic probes and primers for
CC detecting expression of mg genes or mutations, regulatory defects, in
CC this gene, or for isolation of related sequences; and (iv) in (cell-
CC based) gene therapy. (II) are used to raise specific antibodies (Ab); to
CC identify other (extra)cellular products involved in weight regulation,
CC and to screen for agents that disrupt interaction between (II) and other
CC macromolecules. The Ab are used to detect abnormal levels (or function)
CC of (II) (for diagnosis, prognosis or monitoring of treatment); to
CC evaluate (II)-expressing cells intended for cell therapy, and as
CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the
CC mg polypeptide) are used to identify agents (A) that modulate mg
CC activity. (A) are potentially useful for the treatment of body weight
CC disorders, particularly obesity, cachexia or anorexia, or other
CC conditions associated with the mg gene such as hyperpigmentation,
CC hyperphagia and disorders that result in increased metabolic rate
XX
SQ Sequence 6373 BP; 1736 A; 1468 C; 1460 G; 1699 T; 0 U; 10 Other;

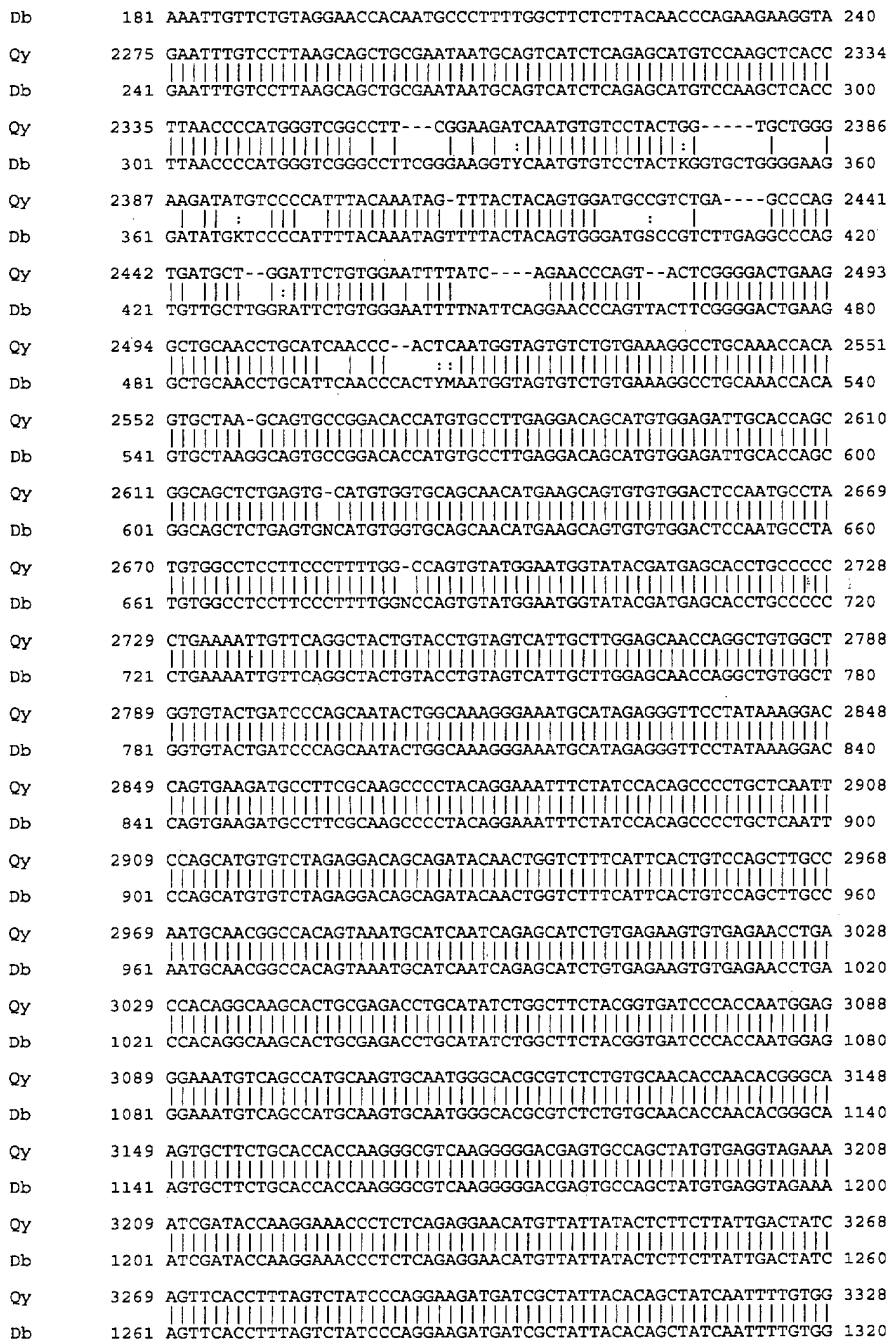
Query Match 37.3%; Score 1341.6; DB 3; Length 6373;
Best Local Similarity 95.2%; Pred. No. 0;
Matches 1498; Conservative 7; Mismatches 42; Indels 26; Gaps 11;

```
Qy       2035 ATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAG 2094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAG 60

Qy       2095 TGGGAGCCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGG 2154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGAGCCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGG 120

Qy       2155 CATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCT 2214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCT 180

Qy       2215 AAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTA 2274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```



Search completed: March 4, 2004, 18:08:57
Job time : 860.6 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2004, 17:06:39 ; Search time 794.925 Seconds
(without alignments)
16549.723 Million cell updates/sec

Title: US-09-787-097-1
Perfect score: 3597
Sequence: 1 atggtggccgcagcggcggc.....aggtgcaaactgaacaatga 3597

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2421054 seqs, 1828716029 residues

Total number of hits satisfying chosen parameters: 4842108

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3494.4	97.1	4072	9	US-09-893-238-16	Sequence 16, Appl
2	3476.8	96.7	8589	9	US-09-893-238-14	Sequence 14, Appl
3	2845.2	79.1	8827	9	US-09-893-238-1	Sequence 1, Appli
4	2028.8	56.4	2625	9	US-09-893-238-18	Sequence 18, Appl
5	1428.8	39.7	2419	9	US-09-893-238-8	Sequence 8, Appli
6	1364.4	37.9	6370	9	US-09-893-238-12	Sequence 12, Appl
7	1200	33.4	6733	16	US-10-197-824-1	Sequence 1, Appli
8	463	12.9	1051	9	US-09-893-238-10	Sequence 10, Appl
c 9	277.4	7.7	625	14	US-10-198-846-5869	Sequence 5869, Ap
10	250.8	7.0	531	9	US-09-864-761-15900	Sequence 15900, A
c 11	246	6.8	246	9	US-09-736-457-966	Sequence 966, App
c 12	246	6.8	246	9	US-09-902-941-966	Sequence 966, App
c 13	246	6.8	246	9	US-09-849-626-966	Sequence 966, App
c 14	246	6.8	246	14	US-10-017-754-966	Sequence 966, App
c 15	246	6.8	246	14	US-10-113-872-966	Sequence 966, App
16	234.4	6.5	207433	15	US-10-277-216-5	Sequence 5, Appli
17	234.4	6.5	207433	16	US-10-126-022-5	Sequence 5, Appli
18	212.8	5.9	415	9	US-09-833-381-1916	Sequence 1916, Ap
19	202.8	5.6	90050	9	US-09-893-238-5	Sequence 5, Appli
20	201.2	5.6	212	9	US-09-796-692-7610	Sequence 7610, Ap
21	201.2	5.6	212	14	US-10-040-862-7610	Sequence 7610, Ap
22	201.2	5.6	212	15	US-10-057-475B-7610	Sequence 7610, Ap
23	201.2	5.6	212	15	US-10-154-884B-7610	Sequence 7610, Ap
24	193	5.4	193	9	US-09-864-761-32404	Sequence 32404, A
c 25	182	5.1	182	9	US-09-864-761-21460	Sequence 21460, A
c 26	174.6	4.9	375	9	US-09-864-761-4720	Sequence 4720, Ap
c 27	169.2	4.7	17056	9	US-09-893-238-3	Sequence 3, Appli
c 28	141.4	3.9	402	9	US-09-864-761-3681	Sequence 3681, Ap
c 29	137	3.8	137	9	US-09-864-761-20449	Sequence 20449, A
c 30	132.2	3.7	226	9	US-09-893-238-6	Sequence 6, Appli
c 31	115.4	3.2	221	9	US-09-893-238-7	Sequence 7, Appli
32	100.2	2.8	5973	9	US-09-893-238-4	Sequence 4, Appli
33	88.4	2.5	629	15	US-10-027-632-270240	Sequence 270240,
34	71.8	2.0	713	13	US-10-040-739-127	Sequence 127, App
35	71.6	2.0	5312	14	US-10-311-455-169	Sequence 169, App
36	66	1.8	594	15	US-10-027-632-264639	Sequence 264639,
37	66	1.8	594	15	US-10-027-632-264640	Sequence 264640,
38	60	1.7	60	10	US-09-908-975-15421	Sequence 15421, A
39	58.6	1.6	65	10	US-09-908-975-29990	Sequence 29990, A
c 40	54.6	1.5	468	15	US-10-027-632-108988	Sequence 108988,
41	48.8	1.4	1731	12	US-10-282-122A-28363	Sequence 28363, A
42	46.2	1.3	2004	12	US-10-282-122A-28310	Sequence 28310, A
c 43	46	1.3	2561	9	US-09-976-740-48	Sequence 48, Appl
c 44	46	1.3	2561	12	US-10-671-242-48	Sequence 48, Appl
c 45	46	1.3	2561	13	US-10-023-529-48	Sequence 48, Appl

ALIGNMENTS

RESULT 1
US-09-893-238-16
; Sequence 16, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 4072
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-893-238-16

Query Match 97.1%; Score 3494.4; DB 9; Length 4072;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 3541; Conservative 0; Mismatches 41; Indels 2; Gaps 2;

```
Qy         15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74
              || |||    |  ||  | ||  |  |  |  | | |   |||| ||   || || ||
Db        234 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAA 293

Qy         75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134
              |    | |  | ||  ||||||||||||||||||||||||||||||||||||||||||||
Db        294 GGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 353
```

```
Qy       135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194
Db       354 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 413

Qy       195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254
Db       414 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 473

Qy       255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314
Db       474 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 533

Qy       315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374
Db       534 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 593

Qy       375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434
Db       594 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 653

Qy       435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494
Db       654 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 713

Qy       495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554
Db       714 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 773

Qy       555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614
Db       774 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 833

Qy       615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 674
Db       834 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 893

Qy       675 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 734
Db       894 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 953

Qy       735 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 794
Db       954 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 1013

Qy       795 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 854
Db      1014 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 1073

Qy       855 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 914
Db      1074 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 1133

Qy       915 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 974
Db      1134 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 1193

Qy       975 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1034
Db      1194 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1253

Qy      1035 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1094
Db      1254 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1313

Qy      1095 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1154
Db      1314 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1373

Qy      1155 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1214
Db      1374 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1433

Qy      1215 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1274
Db      1434 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1493

Qy      1275 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1334
Db      1494 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1553

Qy      1335 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1394
Db      1554 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1613

Qy      1395 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1454
Db      1614 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1673

Qy      1455 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1514
Db      1674 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1733

Qy      1515 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1574
Db      1734 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1793

Qy      1575 CTGGTCAGTGCTTCCCAGACCTGAT-TCCACCATGATGTCAACAGATTTGGCCATTCCAG 1633
Db      1794 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATT-CAG 1852

Qy      1634 CAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCG 1693
Db      1853 CAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCG 1912

Qy      1694 ACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAG 1753
Db      1913 ACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAG 1972

Qy      1754 CAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGG 1813
Db      1973 CAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGG 2032

Qy      1814 CGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTC 1873
Db      2033 CGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTC 2092

Qy      1874 TTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATG 1933
Db      2093 TTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATG 2152

Qy      1934 ACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGA 1993
Db      2153 ACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGA 2212

Qy      1994 TCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGA 2053
Db      2213 TCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGA 2272

Qy      2054 AGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGG 2113
Db      2273 AGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGG 2332

Qy      2114 AGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCAT 2173
Db      2333 AGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCAT 2392

Qy      2174 GTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACC 2233
Db      2393 GTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACC 2452

Qy      2234 ACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGC 2293
Db      2453 ACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGC 2512

Qy      2294 TGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCC 2353
```

```
Db       2513 TGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCC 2572
Qy       2354 TTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTT 2413
Db       2573 TTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTT 2632
Qy       2414 TACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAAC 2473
Db       2633 TACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAAC 2692
Qy       2474 CCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTG 2533
Db       2693 CCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTG 2752
Qy       2534 AAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCAT 2593
Db       2753 AAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCAT 2812
Qy       2594 GTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTG 2653
Db       2813 GTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTG 2872
Qy       2654 TGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGA 2713
Db       2873 TGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGA 2932
Qy       2714 TGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGC 2773
Db       2933 TGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGC 2992
Qy       2774 AACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGG 2833
Db       2993 AACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGG 3052
Qy       2834 GTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCAC 2893
Db       3053 GTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCAC 3112
Qy       2894 AGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTC 2953
Db       3113 AGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTC 3172
Qy       2954 ACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGA 3013
Db       3173 ACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGA 3232
Qy       3014 AGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTG 3073
Db       3233 AGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTG 3292
Qy       3074 ATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCA 3133
Db       3293 ATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCA 3352
Qy       3134 ACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGC 3193
Db       3353 ACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGC 3412
Qy       3194 TATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTC 3253
Db       3413 TATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTC 3472
Qy       3254 TTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAG 3313
Db       3473 TTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAG 3532
Qy       3314 CTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATG 3373
Db       3533 CTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATG 3592
Qy       3374 CCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGG 3433
Db       3593 CCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGG 3652
```

```
Qy       3434 CTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCT 3493
Db       3653 CTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCT 3712
Qy       3494 CTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTA 3553
Db       3713 CTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTA 3772
Qy       3554 ATTTCACCTGGCCCATCAAAATTCAGGTGCAAACTGAACAATGA 3597
Db       3773 ATTTCACCTGGCCCATCAAAATTCAGGTGCAAACTGAACAATGA 3816
```

RESULT 2
US-09-893-238-14
; Sequence 14, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 8589
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-893-238-14

Query Match 96.7%; Score 3476.8; DB 9; Length 8589;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 3524; Conservative 0; Mismatches 42; Indels 2; Gaps 2;

```
Qy         15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74
Db        234 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAA 293
Qy         75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134
Db        294 GGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 353
Qy        135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194
Db        354 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 413
Qy        195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254
Db        414 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 473
Qy        255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314
Db        474 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 533
Qy        315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374
Db        534 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 593
Qy        375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434
Db        594 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 653
Qy        435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494
```

```
Db        654 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 713
Qy        495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554
Db        714 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 773
Qy        555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614
Db        774 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 833
Qy        615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 674
Db        834 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 893
Qy        675 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 734
Db        894 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 953
Qy        735 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 794
Db        954 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 1013
Qy        795 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 854
Db       1014 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 1073
Qy        855 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 914
Db       1074 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 1133
Qy        915 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 974
Db       1134 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 1193
Qy        975 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1034
Db       1194 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1253
Qy       1035 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1094
Db       1254 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1313
Qy       1095 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1154
Db       1314 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1373
Qy       1155 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1214
Db       1374 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1433
Qy       1215 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1274
Db       1434 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1493
Qy       1275 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1334
Db       1494 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1553
Qy       1335 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1394
Db       1554 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1613
Qy       1395 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1454
Db       1614 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1673
Qy       1455 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1514
Db       1674 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1733
Qy       1515 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1574
Db       1734 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1793
Qy       1575 CTGGTCAGTGCTTCCCAGACCTGAT-TCCACCATGATGTCAACAGATTTGGCCATTCCAG 1633
Db       1794 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATT-CAG 1852
Qy       1634 CAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCG 1693
Db       1853 CAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCG 1912
Qy       1694 ACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAG 1753
Db       1913 ACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAG 1972
Qy       1754 CAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGG 1813
Db       1973 CAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGG 2032
Qy       1814 CGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTC 1873
Db       2033 CGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTC 2092
Qy       1874 TTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATG 1933
Db       2093 TTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATG 2152
Qy       1934 ACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGA 1993
Db       2153 ACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGA 2212
Qy       1994 TCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGA 2053
Db       2213 TCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGA 2272
Qy       2054 AGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGG 2113
Db       2273 AGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGG 2332
Qy       2114 AGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCAT 2173
Db       2333 AGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCAT 2392
Qy       2174 GTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACC 2233
Db       2393 GTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACC 2452
Qy       2234 ACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGC 2293
Db       2453 ACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGC 2512
Qy       2294 TGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCC 2353
Db       2513 TGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCC 2572
Qy       2354 TTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTT 2413
Db       2573 TTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTT 2632
Qy       2414 TACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAAATTTTATCAGAAC 2473
Db       2633 TACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAAATTTTATCAGAAC 2692
Qy       2474 CCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTG 2533
Db       2693 CCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTG 2752
Qy       2534 AAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCAT 2593
Db       2753 AAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCAT 2812
Qy       2594 GTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTG 2653
Db       2813 GTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTG 2872
```

```
Qy       2654 TGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGA 2713
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2873 TGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGA 2932

Qy       2714 TGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGC 2773
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2933 TGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGC 2992

Qy       2774 AACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGG 2833
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2993 AACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGG 3052

Qy       2834 GTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCAC 2893
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3053 GTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCAC 3112

Qy       2894 AGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTC 2953
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3113 AGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTC 3172

Qy       2954 ACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGA 3013
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3173 ACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGA 3232

Qy       3014 AGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTG 3073
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3233 AGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTG 3292

Qy       3074 ATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCA 3133
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3293 ATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCA 3352

Qy       3134 ACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGC 3193
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3353 ACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGC 3412

Qy       3194 TATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTC 3253
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3413 TATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTC 3472

Qy       3254 TTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAG 3313
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3473 TTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAG 3532

Qy       3314 CTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATG 3373
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3533 CTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATG 3592

Qy       3374 CCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGG 3433
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3593 CCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGG 3652

Qy       3434 CTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCT 3493
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3653 CTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCT 3712

Qy       3494 CTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTA 3553
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3713 CTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTA 3772

Qy       3554 ATTTCACCTGGCCCATCAAAATTCAGGT 3581
             |||||||||||||||||||||||||||  |
Db       3773 ATTTCACCTGGCCCATCAAAATTCAGAT 3800
```

RESULT 3
US-09-893-238-1
; Sequence 1, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 8827
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-893-238-1

Query Match 79.1%; Score 2845.2; DB 9; Length 8827;
Best Local Similarity 90.3%; Pred. No. 0;
Matches 3064; Conservative 0; Mismatches 328; Indels 2; Gaps 2;

```
Qy        189 ACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAA 248
              |||||||| |||||||| ||||| |||||||||||||| |||||||||| || |||||
Db        677 ACTAACTGGCTCTTCTGGATTTGTAACAGATGGACCTGGGAATTATAAATATAAGACGAA 736

Qy        249 GTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTT 308
              |||||| |||||||||||||||||||||||||||||||||||||| ||||| |||||
Db        737 GTGCACATGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGCTTCAACCATTT 796

Qy        309 TGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACC 368
              ||||||||| ||||| |||||||||||||||||||||||||||||||||| || |||||
Db        797 TGCTACAGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATCTACGCACC 856

Qy        369 GCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCC 428
              || |||||||| |||||||||||||||||||||||| |||||||||||||||||| | ||
Db        857 TCTGATTGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCTCC 916

Qy        429 TGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAA 488
              |||||| || ||| |||||||||| |||||||||||| |||||||||||||||||||||
Db        917 TGAGGTCACTGTCACTTCAGGTTATGCACTGCTGCATTTTTTCAGTGATGCTGCTTATAA 976

Qy        489 TTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCG 548
              | |||||||||||||| |||||| ||||| |||||||| |||| ||||||||||||
Db        977 TCTGACTGGATTTAATATCACTTACAATTTTGACATGTGTCCGAATAATTGCTCAGGCCG 1036

Qy        549 AGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTG 608
              |||||||||||||| |||||| ||||||  ||||||| ||||||||||||||||||||
Db        1037 AGGAGAGTGTAAGAGCAGTAACAGCAGCAGCGCTGTTGAGTGTGAATGTTCTGAAAACTG 1096

Qy        609 GAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGG 668
              |||||| ||  | |||||||||||||||||||||||||||||||||| ||||||| |||||
Db        1097 GAAAGGGGAGTCGTGTGACATTCCTCACTGTACAGACAACTGTGGCTTTCCTCACCGAGG 1156

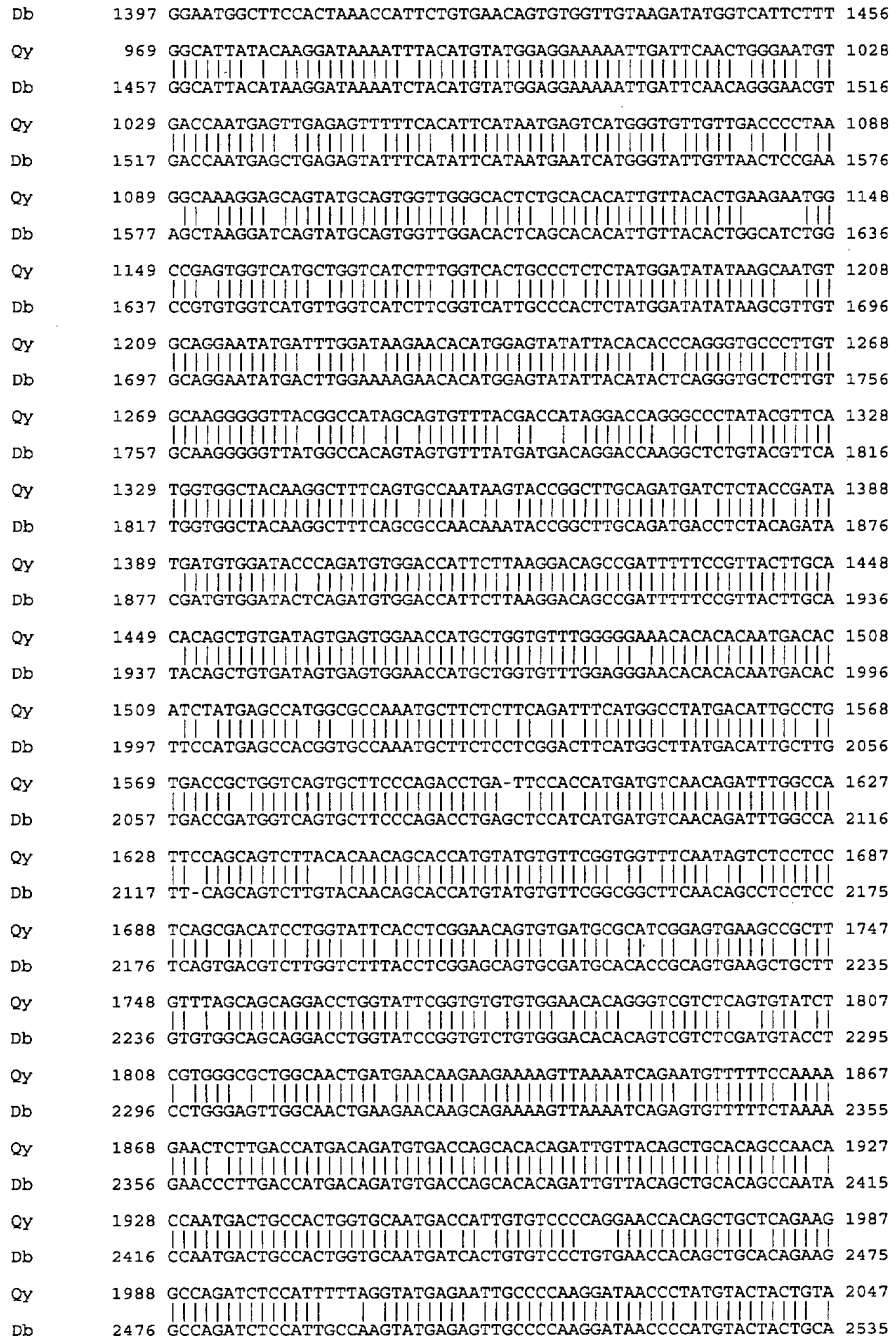
Qy        669 CATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGG 728
              |||||| ||| |||| ||| |||||| |||||||||| | ||||||||||||||||||
Db        1157 CATCTGTAATGCAAGCGATACCAGAGGGTGCTCCTGCTTTCCTCACTGGCAGGGTCCTGG 1216

Qy        729 ATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTT 788
              ||||||| |||||| |||||||||||||| |||||||||||||| |||||||| | ||
Db        1217 ATGTTCAATTCCTGTGCCAGCTAACCAGTCTTTTTGGACTCGAGAAGAATATTCTGATTT 1276

Qy        789 AAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGG 848
              |||||| |||||||| |||||||||||||||||||||||| || ||||||||||||||
Db        1277 AAAGCTTCCCAGAGCCTCTCATAAAGCTGTGGTCAATGGAAATATAATGTGGGTTGTTGG 1336

Qy        849 AGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAG 908
              |||||||||||||||| |||||||| | ||||||| |||||||||||||  |||||||
Db        1337 CGGATATATGTTCAACCATTCAGATTACAGCATGGTTTTAGCGTATGACCTGACTTCTAG 1396

Qy        909 GGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTT 968
              ||| |||||||||||||| |||||||||| |||||||| ||||||||||||||||||
```


Db 1397 GGAATGGCTTCCACTAAACCATTCTGTGAACAGTGTGGTTGTAAGATATGGTCATTCTTT 1456
Qy 969 GGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGT 1028
Db 1457 GGCATTACATAAGGATAAAATCTACATGTATGGAGGAAAAATTGATTCAACAGGGAACGT 1516
Qy 1029 GACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAA 1088
Db 1517 GACCAATGAGCTGAGAGTATTTCATATTCATAATGAATCATGGGTATTGTTAACTCCGAA 1576
Qy 1089 GGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGG 1148
Db 1577 AGCTAAGGATCAGTATGCAGTGGTTGGACACTCAGCACACATTGTTACACTGGCATCTGG 1636
Qy 1149 CCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGT 1208
Db 1637 CCGTGTGGTCATGTTGGTCATCTTCGGTCATTGCCCACTCTATGGATATATAAGCGTTGT 1696
Qy 1209 GCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGT 1268
Db 1697 GCAGGAATATGACTTGGAAAAGAACACATGGAGTATATTACATACTCAGGGTGCTCTTGT 1756
Qy 1269 GCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCA 1328
Db 1757 GCAAGGGGGTTATGGCCACAGTAGTGTTTATGATGACAGGACCAAGGCTCTGTACGTTCA 1816
Qy 1329 TGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATA 1388
Db 1817 TGGTGGCTACAAGGCTTTCAGCGCCAACAAATACCGGCTTGCAGATGACCTCTACAGATA 1876
Qy 1389 TGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCA 1448
Db 1877 CGATGTGGATACTCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCA 1936
Qy 1449 CACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACAC 1508
Db 1937 TACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGAGGGAACACACACAATGACAC 1996
Qy 1509 ATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTG 1568
Db 1997 TTCCATGAGCCACGGTGCCAAATGCTTCTCCTCGGACTTCATGGCTTATGACATTGCTTG 2056
Qy 1569 TGACCGCTGGTCAGTGCTTCCCAGACCTGA-TTCCACCATGATGTCAACAGATTTGGCCA 1627
Db 2057 TGACCGATGGTCAGTGCTTCCCAGACCTGAGCTCCATCATGATGTCAACAGATTTGGCCA 2116
Qy 1628 TTCCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCC 1687
Db 2117 TT-CAGCAGTCTTGTACAACAGCACCATGTATGTGTTCGGCGGCTTCAACAGCCTCCTCC 2175
Qy 1688 TCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTT 1747
Db 2176 TCAGTGACGTCTTGGTCTTTACCTCGGAGCAGTGCGATGCACACCGCAGTGAAGCTGCTT 2235
Qy 1748 GTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCT 1807
Db 2236 GTGTGGCAGCAGGACCTGGTATCCGGTGTCTGTGGGACACACAGTCGTCTCGATGTACCT 2295
Qy 1808 CGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAA 1867
Db 2296 CCTGGGAGTTGGCAACTGAAGAACAAGCAGAAAAGTTAAAATCAGAGTGTTTTTCTAAAA 2355
Qy 1868 GAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACA 1927
Db 2356 GAACCCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAATA 2415
Qy 1928 CCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAG 1987
Db 2416 CCAATGACTGCCACTGGTGCAATGATCACTGTGTCCCTGTGAACCACAGCTGCACAGAAG 2475
Qy 1988 GCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTA 2047
Db 2476 GCCAGATCTCCATTGCCAAGTATGAGAGTTGCCCCAAGGATAACCCCATGTACTACTGCA 2535
Qy 2048 ACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGA 2107
Db 2536 ATAAGAAAACCAGCTGCAGGAGCTGTGCCCTAGACCAGAACTGCCAGTGGGAGCCCCGGA 2595
Qy 2108 ATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAA 2167
Db 2596 ATCAAGAGTGCATCGCCCTGCCGGAAAATATCTGTGGCAATGGCTGGCATTTGGTTGGAA 2655
Qy 2168 ACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTA 2227
Db 2656 ACTCGTGTCTGAAAATCACTACTGCTAAGGAGAATTATGACAATGCTAAATTGTCCTGTA 2715
Qy 2228 GGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTA 2287
Db 2716 GGAACCACAATGCCTTTTTGGCTTCCCTCACATCCCAGAAGAAGGTGGAGTTTGTCCTTA 2775
Qy 2288 AGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGG 2347
Db 2776 AGCAGCTTCGATTAATGCAATCATCTCAAAGTATGTCCAAGCTCACTCTGACTCCATGGG 2835
Qy 2348 TCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAA 2407
Db 2836 TTGGTCTTCGGAAGATCAATGTGTCTTACTGGTGCTGGGAGGATATGTCTCCATTCACAA 2895
Qy 2408 ATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTAT 2467
Db 2896 ATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGTGATGCTGGCTTCTGTGGGATCTTGT 2955
Qy 2468 CAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTG 2527
Db 2956 CAGAGCCTAGTACTCGGGGATTAAAGGCTGCAACCTGCATCAACCCTCTCAATGGCAGCG 3015
Qy 2528 TCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGA 2587
Db 3016 TCTGTGAAAGGCCTGCAAACCACAGTGCCAAGCAGTGCCGGACACCATGTGCCCTGCGGA 3075
Qy 2588 CAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGC 2647
Db 3076 CAGCGTGTGGCGAGTGCACTAGCAGCAGCTCGGAGTGCATGTGGTGCAGTAACATGAAGC 3135
Qy 2648 AGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGT 2707
Db 3136 AGTGTGTGGACTCCAATGCCTACGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGT 3195
Qy 2708 ATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCT 2767
Db 3196 ATACGATGAGCAGCTGCCCACCTGAAAATTGCTCTGGCTACTGTACCTGCAGCCATTGCT 3255
Qy 2768 TGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCA 2827
Db 3256 TGGAGCAGCCAGGCTGTGGTTGGTGTACTGATCCTAGCAATACTGGGAAAGGAAAATGTA 3315
Qy 2828 TAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCT 2887
Db 3316 TTGAGGGCAGCTATAAAGGACCTGTGAAGATGCCGTCACAGGCCTCTGCAGGAAATGTGT 3375
Qy 2888 ATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTT 2947
Db 3376 ATCCACAGCCCCTTCTGAACTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTT 3435
Qy 2948 TCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCT 3007
Db 3436 TCATTCACTGTCCAGCTTGCCAGTGCAACGGACACAGCAAATGCATCAACCAGAGTATCT 3495
Qy 3008 GTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCT 3067
Db 3496 GTGAGAAGTGTGAGGACCTGACCACGGGCAAGCACTGCGAGACCTGCATATCTGGCTTCT 3555
Qy 3068 ACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTC 3127
Db 3556 ATGGTGACCCGACTAATGGAGGCAAATGTCAGCCATGCAAGTGCAATGGGCACGCATCAC 3615

```
Qy       3128 TGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGT 3187
             ||||||||||||||| ||||||||||||| |||||||| || ||||||||||||||||
Db       3616 TGTGCAACACCAACACCGGCAAGTGCTTCTGTACCACCAAAGGTGTCAAGGGGGACGAGT 3675

Qy       3188 GCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATT 3247
             ||||||||||||||||||||||||||||||||||||||||||||| ||||||||| || |
Db       3676 GCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAAAGGAACATGCTACT 3735

Qy       3248 ATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATT 3307
             |||| ||||| |||||||||||||||||||||||| || |||||||||| || ||||| |
Db       3736 ATACCCTTCTCATTGACTATCAGTTCACCTTTAGCCTGTCCCAGGAAGACGACCGCTACT 3795

Qy       3308 ACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCA 3367
             ||||||| ||||| |||||||||||||||||| |||||||||||||||||||||||||||
Db       3796 ACACAGCCATCAACTTTGTGGCTACTCCTGATGAACAAAACAGGGATTTGGACATGTTCA 3855

Qy       3368 TCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAA 3427
             |||||||||||| || |||||||||||||||||||||||  |||| ||| |||| ||||
Db       3856 TCAATGCCTCCAAAAACTTCAACCTCAACATCACCTGGGCCACCAGCTTCCCAGCCGGAA 3915

Qy       3428 CCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATA 3487
             ||||| |||||||||| |||||||||||||||||||||||| ||||| ||||||||||
Db       3916 CCCAGACTGGAGAAGAGGTGCCTGTTGTTTCAAAAACCAACATCAAGGAATACAAAGATA 3975

Qy       3488 GTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATG 3547
             | ||||||||||||||| |||||||||||||||||| ||||| ||||||||||||||||||
Db       3976 GCTTCTCTAATGAGAAATTTGATTTTCGCAACCATCCAAACATCACTTTCTTTGTTTATG 4035

Qy       3548 TCAGTAATTTCACCTGGCCCATCAAAATTCAGGT 3581
             |||||||||||| |||||||||||||||| |
Db       4036 TCAGTAATTTCACTTGGCCCATCAAAATTCAGAT 4069
```

RESULT 4
US-09-893-238-18
; Sequence 18, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 2625
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-893-238-18

Query Match 56.4%; Score 2028.8; DB 9; Length 2625;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 2088; Conservative 0; Mismatches 62; Indels 2; Gaps 2;

```
Qy         15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74
              || |||   |  ||  | ||  |  |  | | |  |||| ||   || || ||
Db        234 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAA 293

Qy         75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134
              |    | |  | ||  ||||||||||||||||||||||||||||||||||||||||||
Db        294 GGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 353

Qy        135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        354 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 413

Qy        195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 473

Qy        255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 533

Qy        315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 593

Qy        375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 653

Qy        435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        654 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 713

Qy        495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        714 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 773

Qy        555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        774 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 833

Qy        615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        834 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 893

Qy        675 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        894 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 953

Qy        735 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        954 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 1013

Qy        795 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1014 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 1073

Qy        855 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1074 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 1133

Qy        915 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1134 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 1193

Qy        975 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1194 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1253

Qy       1035 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1254 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1313

Qy       1095 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1314 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1373

Qy       1155 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1374 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1433

Qy       1215 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1434 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1493
```

```
Qy       1275 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1334
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1494 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1553

Qy       1335 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1394
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1554 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1613

Qy       1395 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1454
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1614 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1673

Qy       1455 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1514
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1674 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1733

Qy       1515 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1574
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1734 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1793

Qy       1575 CTGGTCAGTGCTTCCCAGACCTGAT-TCCACCATGATGTCAACAGATTTGGCCATTCCAG 1633
             ||||||||||||||||||||||||| ||||||||||||||||||||||||||| |||
Db       1794 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATT-CAG 1852

Qy       1634 CAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCG 1693
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1853 CAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCG 1912

Qy       1694 ACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAG 1753
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1913 ACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAG 1972

Qy       1754 CAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGG 1813
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1973 CAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGG 2032

Qy       1814 CGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTC 1873
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2033 CGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTC 2092

Qy       1874 TTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATG 1933
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2093 TTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATG 2152

Qy       1934 ACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGA 1993
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2153 ACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGA 2212

Qy       1994 TCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGA 2053
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2213 TCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGA 2272

Qy       2054 AGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGG 2113
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2273 AGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGG 2332

Qy       2114 AGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGG 2165
             |||||||||||||||||  |   ||  |  |  ||  |  |  ||  ||
Db       2333 AGTGCATTGCCCTGCCCGGTAGGCCTTGCAGGGTCATCTTGGTGTGTGTGGG 2384
```

RESULT 5
US-09-893-238-8
; Sequence 8, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 2419
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-893-238-8

Query Match 39.7%; Score 1428.8; DB 9; Length 2419;
Best Local Similarity 89.4%; Pred. No. 0;
Matches 1560; Conservative 0; Mismatches 182; Indels 2; Gaps 2;

```
Qy        189 ACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAA 248
              |||||||| |||||||| ||||| |||||||||||||| |||||||||| || |||||
Db        677 ACTAACTGGCTCTTCTGGATTTGTAACAGATGGACCTGGGAATTATAAATATAAGACGAA 736

Qy        249 GTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTT 308
              |||||| |||||||||||||||||||||||||||||||||||||||||| ||||| |||||
Db        737 GTGCACATGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGCTTCAACCATTT 796

Qy        309 TGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACC 368
              |||||||| ||||| |||||||||||||||||||||||||||||||||||||| || |||||
Db        797 TGCTACAGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATCTACGCACC 856

Qy        369 GCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCC 428
              ||  ||||||| |||||||||||||||||||||||| ||||||||||||||||| |  ||
Db        857 TCTGATTGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCTCC 916

Qy        429 TGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAA 488
              ||||||   || ||| |||||||||| |||||||||||| ||||||||||||||||||
Db        917 TGAGGTCACTGTCACTTCAGGTTATGCACTGCTGCATTTTTTCAGTGATGCTGCTTATAA 976

Qy        489 TTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCG 548
              | |||||||||||||| |||||| |||||| |||||| |||| ||||||||||||||||
Db        977 TCTGACTGGATTTAATATCACTTACAATTTTGACATGTGTCCGAATAATTGCTCAGGCCG 1036

Qy        549 AGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTG 608
              |||||||||||||| |||||| |||||   ||||||| |||||||||||||||||||||
Db        1037 AGGAGAGTGTAAGAGCAGTAACAGCAGCAGCGCTGTTGAGTGTGAATGTTCTGAAAACTG 1096

Qy        609 GAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGG 668
              |||||| ||  | |||||||||||||||||||||||||||||||||||||| |||||||
Db        1097 GAAAGGGGAGTCGTGTGACATTCCTCACTGTACAGACAACTGTGGCTTTCCTCACCGAGG 1156

Qy        669 CATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGG 728
              |||||| ||| |||| ||| |||||| |||||||||| | ||||||||||||||||||
Db        1157 CATCTGTAATGCAAGCGATACCAGAGGGTGCTCCTGCTTTCCTCACTGGCAGGGTCCTGG 1216

Qy        729 ATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTT 788
              ||||||| ||||||| |||||||||||||| |||||||||||||| ||||||||| | ||
Db        1217 ATGTTCAATTCCTGTGCCAGCTAACCAGTCTTTTTGGACTCGAGAAGAATATTCTGATTT 1276

Qy        789 AAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGG 848
              |||||| |||||||| |||||||||||||||||||||||||||||| || ||||||||||
Db        1277 AAAGCTTCCCAGAGCCTCTCATAAAGCTGTGGTCAATGGAAATATAATGTGGGTTGTTGG 1336

Qy        849 AGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAG 908
              ||||||||||||||||| |||||||| | ||||||| |||||||||||||  |||||||
Db        1337 CGGATATATGTTCAACCATTCAGATTACAGCATGGTTTTAGCGTATGACCTGACTTCTAG 1396

Qy        909 GGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTT 968
              ||| |||||||||||||| |||||||||| |||||||| |||||||||||||||||||||
Db        1397 GGAATGGCTTCCACTAAACCATTCTGTGAACAGTGTGGTTGTAAGATATGGTCATTCTTT 1456

Qy        969 GGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGT 1028
              ||||||| | |||||||||| |||||||||||||||||||||||||||| ||||| ||
```

```
Db       1457 GGCATTACATAAGGATAAAATCTACATGTATGGAGGAAAAATTGATTCAACAGGGAACGT 1516

Qy       1029 GACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAA 1088
              |||||||||| ||||||| ||||| |||||||||| |||||||| ||||| || || ||
Db       1517 GACCAATGAGCTGAGAGTATTTCATATTCATAATGAATCATGGGTATTGTTAACTCCGAA 1576

Qy       1089 GGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGG 1148
                || ||||| |||||||||||||||||| ||||| |||||||||||||||||||     |||
Db       1577 AGCTAAGGATCAGTATGCAGTGGTTGGACACTCAGCACACATTGTTACACTGGCATCTGG 1636

Qy       1149 CCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGT 1208
              ||| |||||||| |||||||| ||||| |||||||||||||||||||||||||   |||
Db       1637 CCGTGTGGTCATGTTGGTCATCTTCGGTCATTGCCCACTCTATGGATATATAAGCGTTGT 1696

Qy       1209 GCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGT 1268
              |||||||||||| ||||| |||||||||||||||||||||| || |||||||| |||||
Db       1697 GCAGGAATATGACTTGGAAAAGAACACATGGAGTATATTACATACTCAGGGTGCTCTTGT 1756

Qy       1269 GCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCA 1328
              |||||||||||| ||||| || |||||||| ||  | ||||||| ||| || ||||||||
Db       1757 GCAAGGGGGTTATGGCCACAGTAGTGTTTATGATGACAGGACCAAGGCTCTGTACGTTCA 1816

Qy       1329 TGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATA 1388
              |||||||||||||||||||| ||||| || |||||||||||||||| |||||| ||||
Db       1817 TGGTGGCTACAAGGCTTTCAGCGCCAACAAATACCGGCTTGCAGATGACCTCTACAGATA 1876

Qy       1389 TGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCA 1448
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1877 CGATGTGGATACTCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCA 1936

Qy       1449 CACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACAC 1508
              |||||||||||||||||||||||||||||||||||| || ||||||||||||||||||
Db       1937 TACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGAGGGAACACACACAATGACAC 1996

Qy       1509 ATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTG 1568
              || |||||||| || |||||||||||||| || || |||||||| |||||||||| ||
Db       1997 TTCCATGAGCCACGGTGCCAAATGCTTCTCCTCGGACTTCATGGCTTATGACATTGCTTG 2056

Qy       1569 TGACCGCTGGTCAGTGCTTCCCAGACCTGA-TTCCACCATGATGTCAACAGATTTGGCCA 1627
              |||||| |||||||||||||||||||||||  |||| |||||||||||||||||||||||
Db       2057 TGACCGATGGTCAGTGCTTCCCAGACCTGAGCTCCATCATGATGTCAACAGATTTGGCCA 2116

Qy       1628 TTCCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCC 1687
              || |||||||||  |||||||||||||||||||||||||||||| || |||||| ||||||||
Db       2117 TT-CAGCAGTCTTGTACAACAGCACCATGTATGTGTTCGGCGGCTTCAACAGCCTCCTCC 2175

Qy       1688 TCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTT 1747
              |||| ||| || |||| || |||||||| ||||| ||||| || || |||||||| ||||
Db       2176 TCAGTGACGTCTTGGTCTTTACCTCGGAGCAGTGCGATGCACACCGCAGTGAAGCTGCTT 2235

Qy       1748 GTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCT 1807
              || | |||||||||||||||| |||||| |||| |||||   |||||||  |||| ||
Db       2236 GTGTGGCAGCAGGACCTGGTATCCGGTGTCTGTGGGACACACAGTCGTCTCGATGTACCT 2295

Qy       1808 CGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAA 1867
              | |||| | |||||||||| ||||||| |||||||||||||||||| ||||||| ||||
Db       2296 CCTGGGAGTTGGCAACTGAAGAACAAGCAGAAAAGTTAAAATCAGAGTGTTTTTCTAAAA 2355

Qy       1868 GAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACA 1927
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       2356 GAACCCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAATA 2415

Qy       1928 CCAA 1931
              ||||
Db       2416 CCAA 2419
```

```
RESULT 6
US-09-893-238-12
; Sequence 12, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT:  Nagle, D.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
;  FILE REFERENCE: 7853-237
;  CURRENT APPLICATION NUMBER: US/09/893,238
;  CURRENT FILING DATE:  2001-06-27
;  PRIOR APPLICATION NUMBER: 09/245,041
;  PRIOR FILING DATE: 1999-02-05
;  PRIOR APPLICATION NUMBER: 60/093,630
;  PRIOR FILING DATE: 1998-07-21
;  PRIOR APPLICATION NUMBER: 60/104,978
;  PRIOR FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 12
;   LENGTH: 6370
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-893-238-12

  Query Match             37.9%;  Score 1364.4;  DB 9;  Length 6370;
  Best Local Similarity   95.0%;  Pred. No. 0;
  Matches 1492;  Conservative    7;  Mismatches   48;  Indels   23;  Gaps    8;

Qy       2035 ATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAG 2094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAG 60

Qy       2095 TGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGG 2154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGG 120

Qy       2155 CATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCT 2214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 CATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCT 180

Qy       2215 AAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTA 2274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 AAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTA 240

Qy       2275 GAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACC 2334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 GAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACC 300

Qy       2335 TTAACCCCATGGGTCGGCCTT---CGGAAGATCAATGTGTCCTACTGG-----TGCTGGG 2386
              |||||||||||||||| | |    | | |:|||||||||||||:|      |   |
Db         301 TTAACCCCATGGGTCGGGCCTTCGGGAAGGTYCAATGTGTCCTACTKGGTGCTGGGGAAG 360

Qy       2387 AAGATATGTCCCCATTTACAAATAG-TTTACTACAGTGGATGCCGTCTG----AGCCCAG 2441
              | || :  |||  |||||||||| |||||||||||    :   |      ||||||
Db         361 GATATGKTCCCCATTTTACAAATAGTTTTACTACAGTGGGATGSCCGTCTTGAGGCCCAG 420

Qy       2442 TGATGCT--GGATTCTGTGGAATTTT-----ATCAGAACCCAGTACTCGGGGACTGAAGG 2494
              || |||   |:|||||||| | |||      |  |   |       ||||||||||||||
Db         421 TGTTGCTTGGRATTCTGTGGGAATTTTATTCAGGAACCCAGTTACTTCGGGGACTGAAGG 480

Qy       2495 CTGCAACCTGCAT--CAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAG 2552
              |||||||||||||   | ||   ::|||||||||||||||||||||||||||||||||
Db         481 CTGCAACCTGCATTCAACCCACTYMAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAG 540

Qy       2553 TGCTAA-GCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCG 2611
              |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 TGCTAAGGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCG 600

Qy       2612 GCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATG 2671
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 GCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATG 660

Qy       2672 TGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTG 2731
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 TGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTG 720

Qy       2732 AAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGT 2791
```

```
Db        721 AAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGT 780

Qy       2792 GTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAG 2851

Db        781 GTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAG 840

Qy       2852 TGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCA 2911

Db        841 TGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCA 900

Qy       2912 GCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAAT 2971

Db        901 GCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAAT 960

Qy       2972 GCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCA 3031

Db        961 GCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCA 1020

Qy       3032 CAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGA 3091

Db       1021 CAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGA 1080

Qy       3092 AATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGT 3151

Db       1081 AATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGT 1140

Qy       3152 GCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATC 3211

Db       1141 GCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATC 1200

Qy       3212 GATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGT 3271

Db       1201 GATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGT 1260

Qy       3272 TCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTA 3331

Db       1261 TCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTA 1320

Qy       3332 CTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACC 3391

Db       1321 CTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACC 1380

Qy       3392 TCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTG 3451

Db       1381 TCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTG 1440

Qy       3452 TTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATT 3511

Db       1441 TTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATT 1500

Qy       3512 TTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCA 3571

Db       1501 TTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCA 1560

Qy       3572 AAATTCAGGT 3581

Db       1561 AAATTCAGAT 1570


RESULT 7
US-10-197-824-1
; Sequence 1, Application US/10197824
; Publication No. US20040023219A1
; GENERAL INFORMATION:
;  APPLICANT: ORIGENE TECHNOLOGIES INC
;  TITLE OF INVENTION: NOVEL PROSTATE CANCER GENES
;  FILE REFERENCE: 9U 106 R1
;  CURRENT APPLICATION NUMBER: US/10/197,824
;  CURRENT FILING DATE:  2002-07-19
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;  LENGTH: 6733
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY: CDS
;  LOCATION: (575)..(4711)
;  OTHER INFORMATION:
US-10-197-824-1

  Query Match             33.4%;  Score 1200;  DB 16;  Length 6733;
  Best Local Similarity   62.0%;  Pred. No. 0;
  Matches 2179;  Conservative    0;  Mismatches 1245;  Indels   92;  Gaps   14;

Qy        110 GCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGC 169

Db        789 GCCGCTGTGTCAACTCCACCTGCCTCTGCGACCCGGGCTGGGTGGGGGACCAGTGCCAGC 848

Qy        170 ACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAA 229

Db        849 ACTGCCAGGGCAGGTTCAAGTTAACAGAACCTTCTGGATATTTAACAGATGGCCCAATTA 908

Qy        230 ATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGA 289

Db        909 ACTATAAATATAAAACTAAATGTACTTGGCTCATTGAAGGCTATCCAAATGCAGTGTTAA 968

Qy        290 GACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATG 349

Db        969 GATTAAGATTCAATCATTTTGCTACAGAATGTAGCTGGGATCATATGTATGTTTATGATG 1028

Qy        350 GGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAG 409

Db       1029 GAGATTCAATATATGCACCTTTAATAGCTGTACTTAGTGGTTTGATAGTCCCTGAAATAA 1088

Qy        410 ATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTT 469

Db       1089 GGGGCAATGAAACTGTGCCTGAAGTTGTTACTACATCTGGCTATGCACTGTTACATTTTT 1148

Qy        470 TTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTC 529

Db       1149 TTAGTGATGCTGCGTATAATCTAACTGGTTTCAACATTTTCTATTCAATCAATTCTTGTC 1208

Qy        530 CAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGT---AATAGCAGCGATACTGTTG 586

Db       1209 CTAACAATTGCTCTGGTCATGGGAAGTGTACAACTAGTGTCTCTGTTCCAAGTCAAGTAT 1268

Qy        587 AATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACA 646

Db       1269 ATTGTGAATGTGATAAATACTGGAAGGGTGAAGCTTGTGATATTCCTTACTGTAAAGCCA 1328

Qy        647 ACTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCT 706

Db       1329 ATTGCGGCAGTCCAGATCACGGTTACTGTGACCTGACTGGAGAAAAATTATGTGTCTGCA 1388

Qy        707 TCTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGA 766

Db       1389 ATGATAGTTGGCAAGGTCCTGATTGTTCTTTGAATGTTCCCTCTACTGAGTCTTACTGGA 1448

Qy        767 CTCGAGAGGAATATTCTAACTTAAAGC------TCCCCAGAGCATCTCATAAAGCTGTGG 820

Db       1449 TTCTGCCAAACGTTAAACCCTTCAGTCCTTCTGTAGGTCGGGCTTCACATAAAGCAGTTT 1508

Qy        821 TCAATGGAAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACA 880

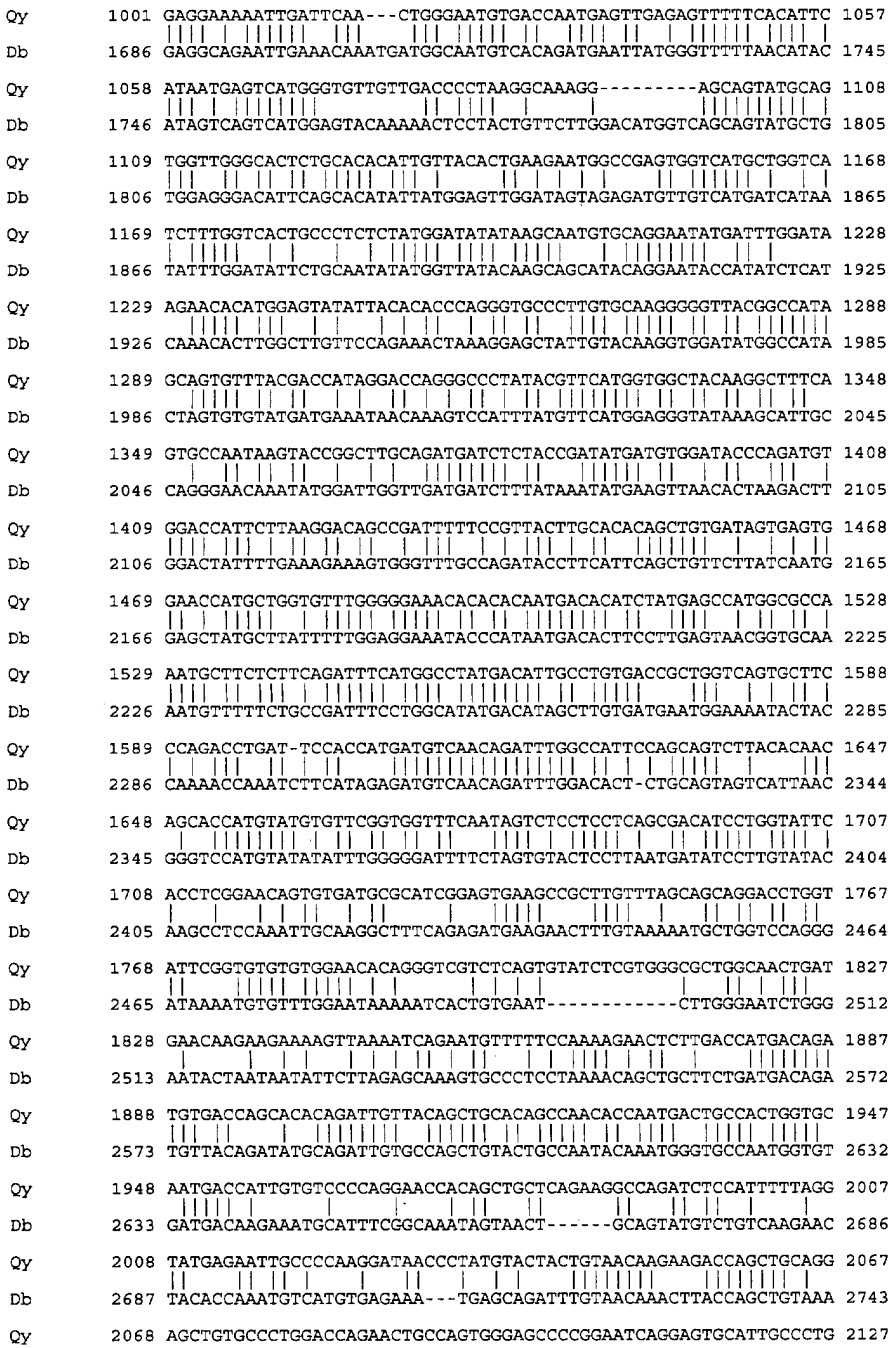
Db       1509 TACACGGGAAATTTATGTGGGTGATTGGTGGATATACTTTTAACTACAGTTCTTTTCAAA 1568

Qy        881 TGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACA 940

Db       1569 TGGTCCTAAATTACAATTTAGAAAGCAGTATATGGAATGTAGGAACTCCATCAAGGGGAC 1628

Qy        941 ATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACATGTATG 1000

Db       1629 CT---CTCCAGAGATATGGACACTCTCTTGCTTTATATCAGGAAAACATCTTTATGTATG 1685
```


Qy 1001 GAGGAAAAATTGATTCAA---CTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTC 1057
Db 1686 GAGGCAGAATTGAAACAAATGATGGCAATGTCACAGATGAATTATGGGTTTTTAACATAC 1745
Qy 1058 ATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGG---------AGCAGTATGCAG 1108
Db 1746 ATAGTCAGTCATGGAGTACAAAAACTCCTACTGTTCTTGGACATGGTCAGCAGTATGCTG 1805
Qy 1109 TGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCA 1168
Db 1806 TGGAGGGACATTCAGCACATATTATGGAGTTGGATAGTAGAGATGTTGTCATGATCATAA 1865
Qy 1169 TCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATA 1228
Db 1866 TATTTGGATATTCTGCAATATATGGTTATACAAGCAGCATACAGGAATACCATATCTCAT 1925
Qy 1229 AGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATA 1288
Db 1926 CAAACACTTGGCTTGTTCCAGAAACTAAAGGAGCTATTGTACAAGGTGGATATGGCCATA 1985
Qy 1289 GCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCA 1348
Db 1986 CTAGTGTGTATGATGAAATAACAAAGTCCATTTATGTTCATGGAGGGTATAAAGCATTGC 2045
Qy 1349 GTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGT 1408
Db 2046 CAGGGAACAAATATGGATTGGTTGATGATCTTTATAAATATGAAGTTAACACTAAGACTT 2105
Qy 1409 GGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTG 1468
Db 2106 GGACTATTTTGAAAGAAAGTGGGTTTGCCAGATACCTTCATTCAGCTGTTCTTATCAATG 2165
Qy 1469 GAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGCCATGGCGCCA 1528
Db 2166 GAGCTATGCTTATTTTTGGAGGAAATACCCATAATGACACTTCCTTGAGTAACGGTGCAA 2225
Qy 1529 AATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTC 1588
Db 2226 AATGTTTTTCTGCCGATTTCCTGGCATATGACATAGCTTGTGATGAATGGAAAATACTAC 2285
Qy 1589 CCAGACCTGAT-TCCACCATGATGTCAACAGATTTGGCCATTCCAGCAGTCTTACACAAC 1647
Db 2286 CAAAACCAAATCTTCATAGAGATGTCAACAGATTTGGACACT-CTGCAGTAGTCATTAAC 2344
Qy 1648 AGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTC 1707
Db 2345 GGGTCCATGTATATATTTGGGGGATTTTCTAGTGTACTCCTTAATGATATCCTTGTATAC 2404
Qy 1708 ACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGT 1767
Db 2405 AAGCCTCCAAATTGCAAGGCTTTCAGAGATGAAGAACTTTGTAAAAATGCTGGTCCAGGG 2464
Qy 1768 ATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGAT 1827
Db 2465 ATAAAATGTGTTTGGAATAAAAATCACTGTGAAT------------CTTGGGAATCTGGG 2512
Qy 1828 GAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGA 1887
Db 2513 AATACTAATAATATTCTTAGAGCAAAGTGCCCTCCTAAAACAGCTGCTTCTGATGACAGA 2572
Qy 1888 TGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGC 1947
Db 2573 TGTTACAGATATGCAGATTGTGCCAGCTGTACTGCCAATACAAATGGGTGCCAATGGTGT 2632
Qy 1948 AATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGG 2007
Db 2633 GATGACAAGAAATGCATTTCGGCAAATAGTAACT------GCAGTATGTCTGTCAAGAAC 2686
Qy 2008 TATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGG 2067
Db 2687 TACACCAAATGTCATGTGAGAAA---TGAGCAGATTTGTAACAAACTTACCAGCTGTAAA 2743
Qy 2068 AGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTG 2127
Db 2744 AGCTGTTCACTAAACTTGAATTGCCAGTGGGATCAGAGACAGCAAGAATGCCAGGCTTTA 2803
Qy 2128 CCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACT 2187
Db 2804 CCAGCTCATCTTTGTGGAGAAGGATGGAGTCATATTGGGGATGCTTGTCTTAGAGTCAAT 2863
Qy 2188 ACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTG 2247
Db 2864 TCCAGTAGAGAAAACTATGACAATGCAAAACTTTATTGCTATAATCTTAGTGGAAATCTT 2923
Qy 2248 GCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAG 2307
Db 2924 GCTTCATTAACAACCTCAAAAGAAGTAGAATTTGTTC----------------TGGAT 2965
Qy 2308 TCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAAT 2367
Db 2966 GAAATACAGAAGTATACACAACAGAAAGTATCACCTTGGGTAGGCTTGCGCAAGATCAAT 3025
Qy 2368 GTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATG 2427
Db 3026 ATATCCTATTGGGGATGGGAAGACATGTCTCCTTTTACAAACACAACACTACAGTGGCTT 3085
Qy 2428 CCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGA 2487
Db 3086 CCTGGCGAACCCAATGATTCTGGGTTTTGTGCATATCTGGAAAGGGCTGCAGTGGCAGGC 3145
Qy 2488 CTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTG----- 2542
Db 3146 TTAAAAGCTAATCCTTGTACATCTATGGCAAATGGCCTTGTCTGTGAAAAACCTGTTGTT 3205
Qy 2543 ----CAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGA 2598
Db 3206 AGTCCAAATCAAAATGCGAGGCCGTGCAAAAAGCCATGCTCTCTGAGGACATCATGTTCC 3265
Qy 2599 GATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGAC 2658
Db 3266 AACTGTACAAGCAATGGCATGGAGTGTATGTGGTGCAGCAGTACGAAACGATGTGTTGAC 3325
Qy 2659 TCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGC 2718
Db 3326 TCTAATGCCTATATCATCTCTTTTCCATATGGACAATGTCTAGAGTGGCAAAC---TGCC 3382
Qy 2719 ACCTGCCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCA 2778
Db 3383 ACCTGCTCCCCTCTAAATTGTTCTGGATTGAGAACCTGTGGACAGTGTTTGGAACAGCCT 3442
Qy 2779 GGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCC 2838
Db 3443 GGATGTGGCTGGTGCAATGATCCTAGTAATACAGGAAGAGGACATTGCATTGAAGGTTCT 3502
Qy 2839 TATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCC 2898
Db 3503 TCACGGGGACCAATGAAG----------------CTTATTGGAATGCACCACAATGAGATG 3547
Qy 2899 CTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGT 2958
Db 3548 GTTCTTGACACCAATCTTTGCCCCAAAGAAAAGAACTATGAGTGGTCCTTTATCCAGTGT 3607
Qy 2959 CCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGT 3018
Db 3608 CCAGCTTGCCAGTGTAATGGACATAGCACTTGCATCAATAATAATGTGTGCGAACAGTGT 3667
Qy 3019 GAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCC 3078
Db 3668 AAAAATCTCACCACAGGAAAGCAGTGTCAAGATTGTATGCCAGGTTATTATGGAGATCCA 3727
Qy 3079 ACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACC 3138
Db 3728 ACCAATGGTGGACAGTGCACAGCTTGTACATGCAGTGGCCATGCAAATATCTGTCATCTG 3787
Qy 3139 AACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGT 3198

```
Db       3788 CACACAGGAAAATGTTTCTGCACAACTAAAGGAATAAAAGGTGACCAATGCCAATTATGT 3847

Qy       3199 GAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTT 3258

Db       3848 GACTCTGAAAATCGCTATGTTGGTAATCCACTTAGAGGAACATGTTATTACAGCCTTTTG 3907

Qy       3259 ATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATC 3318

Db       3908 ATTGATTATCAATTTACCTTCAGCTTATTACAGGAAGATGATCGCCACCATACTGCCATA 3967

Qy       3319 AATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCC 3378

Db       3968 AACTTTATAGCAAACCCAGAACAGTCGAACAAAAATCTGGATATATCAATTAATGCATCA 4027

Qy       3379 AAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGA 3438

Db       4028 AACAACTTTAATCTCAACATTTCGTGGTCTGTCGGTTCAACAGCTGGAACAATATCTGGG 4087

Qy       3439 GAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAAT 3498

Db       4088 GAAGAGACTTCTATAGTTTCCAAGAATAATATAAAGGAATACAGAGATAGTTTTTCCTAT 4147

Qy       3499 GAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTC 3558

Db       4148 GAAAAATTTAACTTTAGAAGCAATCCTAACATTACATTCTATGTGTACGTCAGCAACTTT 4207

Qy       3559 ACCTGGCCCATCAAAATTCAGGTGCAAACTGAACAA 3594

Db       4208 TCCTGGCCTATTAAAATACAGATTGCATTCTCACAA 4243


RESULT 8
US-09-893-238-10
; Sequence 10, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT:  Nagle, D.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
;  FILE REFERENCE: 7853-237
;  CURRENT APPLICATION NUMBER: US/09/893,238
;  CURRENT FILING DATE:  2001-06-27
;  PRIOR APPLICATION NUMBER: 09/245,041
;  PRIOR FILING DATE: 1999-02-05
;  PRIOR APPLICATION NUMBER: 60/093,630
;  PRIOR FILING DATE: 1998-07-21
;  PRIOR APPLICATION NUMBER: 60/104,978
;  PRIOR FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10
;   LENGTH: 1051
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-893-238-10

  Query Match             12.9%;  Score 463;  DB 9;  Length 1051;
  Best Local Similarity   85.2%;  Pred. No. 1.3e-132;
  Matches  517;  Conservative    0;  Mismatches   90;  Indels    0;  Gaps    0;

Qy         15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74

Db        441 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAA 500

Qy         75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134

Db        501 GGAATGTGACCGGCCGTGTGTCAACGGCGGCCGCTGCAACCCTGGCACCGGCCAGTGCGT 560

Qy        135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194

Db        561 CTGCCCCACGGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 620

Qy        195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254

Db        621 TGGCTCTTCTGGATTTGTAACAGATGGACCTGGGAATTATAAATATAAGACGAAGTGCAC 680

Qy        255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314

Db        681 ATGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGCTTCAACCATTTTGCTAC 740

Qy        315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374

Db        741 AGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATCTACGCACCTCTGAT 800

Qy        375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434

Db        801 TGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCTCCTGAGGT 860

Qy        435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494

Db        861 CACTGTCACTTCAGGTTATGCACTGCTGCATTTTTTCAGTGATGCTGCTTATAATCTGAC 920

Qy        495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554

Db        921 TGGATTTAATATCACTTACAATTTTGACATGTGTCCGAATAATTGCTCAGGCCGAGGAGA 980

Qy        555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614

Db        981 GTGTAAGAGCAGTAACAGCAGCAGCGCTGTTGAGTGTGAATGTTCTGAAAACTGGAAAGG 1040

Qy        615 TGAAGCA 621

Db       1041 GGCCGGA 1047


RESULT 9
US-10-198-846-5869/c
; Sequence 5869, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
;  APPLICANT: Lillie, James
;  APPLICANT:  Xu, Yongyao
;  APPLICANT:  Wang, Youzhen
;  APPLICANT:  Steinmann, Kathleen
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
;  TITLE OF INVENTION:  FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
;  TITLE OF INVENTION:  THERAPY OF BREAST CANCER
;  FILE REFERENCE: MRI-049
;  CURRENT APPLICATION NUMBER: US/10/198,846
;  CURRENT FILING DATE:  2002-07-18
;  PRIOR APPLICATION NUMBER: 60/306,220
;  PRIOR FILING DATE: 2001-07-18
;  NUMBER OF SEQ ID NOS: 14084
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5869
;   LENGTH: 625
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 3, 4, 5, 6, 7, 8, 10, 12, 14, 15, 17, 26, 28, 29
;   OTHER INFORMATION: n = A,T,C or G
US-10-198-846-5869

  Query Match             7.7%;  Score 277.4;  DB 14;  Length 625;
  Best Local Similarity   99.6%;  Pred. No. 6.4e-75;
  Matches  278;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2477 GTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAA 2536

Db        340 GTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAA 281

Qy       2537 GGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTG 2596

Db        280 GGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTG 221
```

```
Qy       2597 GAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGG 2656
             ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        220 GAGATTGCACCAGCGGCAGCTCTGAGTGTATGTGGTGCAGCAACATGAAGCAGTGTGTGG 161

Qy       2657 ACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGA 2716
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 ACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGA 101

Qy       2717 GCACCTGCCCCCCCTGAAAATTGTTCAGGCTACTGTACCT 2755
             |||||||||||||||||||||||||||||||||||||||
Db        100 GCACCTGCCCCCCCTGAAAATTGTTCAGGCTACTGTACCT 62
```

RESULT 10
US-09-864-761-15900
; Sequence 15900, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 15900
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL132773.14
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
US-09-864-761-15900

Query Match 7.0%; Score 250.8; DB 9; Length 531;
Best Local Similarity 97.3%; Pred. No. 1.1e-66;
Matches 255; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy        973 TTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACC 1032
              || | | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 TTTTCCTAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACC 200

Qy       1033 AATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCA 1092
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 AATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCA 260

Qy       1093 AAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGA 1152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 AAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGA 320

Qy       1153 GTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAG 1212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 GTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAG 380

Qy       1213 GAATATGATTTGGATAAGAACA 1234
              ||||||||||||| || | | |
Db        381 GAATATGATTTGGGTAGGTATA 402
```

RESULT 11
US-09-736-457-966/c
; Sequence 966, Application US/09736457
; Patent No. US20020168637A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Ligun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-966

Query Match 6.8%; Score 246; DB 9; Length 246;
Best Local Similarity 100.0%; Pred. No. 1.9e-65;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2293 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 2352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy       2353 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 2412
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy       2413 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 2472
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67

Qy       2473 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 2532
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy       2533 GAAAGG 2538
                ||||||
Db          6 GAAAGG 1


RESULT 12
US-09-902-941-966/c
; Sequence 966, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
;  APPLICANT: Henderson, Robert A.
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Watanabe, Yoshihiro
;  APPLICANT:  Johnson, Jeffrey C.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Marnerakis, Margarita
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  McNabb, Andria
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.478C17
;  CURRENT APPLICATION NUMBER: US/09/902,941
;  CURRENT FILING DATE:  2001-07-10
;  NUMBER OF SEQ ID NOS: 2002
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 966
;   LENGTH: 246
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-902-941-966

  Query Match             6.8%;  Score 246;  DB 9;  Length 246;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-65;
  Matches  246;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2293 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 2352
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy       2353 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 2412
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy       2413 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAAATTTTATCAGAA 2472
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAAATTTTATCAGAA 67

Qy       2473 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 2532
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy       2533 GAAAGG 2538
                ||||||
Db          6 GAAAGG 1


RESULT 13
US-09-849-626-966/c
; Sequence 966, Application US/09849626
; Publication No. US20020197669A1
; GENERAL INFORMATION:
;  APPLICANT: Bangur, Chaitanya
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Switzer, Anne
;  APPLICANT:  McNeill, Patricia
;  APPLICANT:  Clapper, Jonathan
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.478C16
;  CURRENT APPLICATION NUMBER: US/09/849,626
;  CURRENT FILING DATE:  2001-05-03
;  NUMBER OF SEQ ID NOS: 1926
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 966
;   LENGTH: 246
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-849-626-966

  Query Match             6.8%;  Score 246;  DB 9;  Length 246;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-65;
  Matches  246;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2293 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 2352
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy       2353 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 2412
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy       2413 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAAATTTTATCAGAA 2472
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAAATTTTATCAGAA 67

Qy       2473 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 2532
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy       2533 GAAAGG 2538
                ||||||
Db          6 GAAAGG 1


RESULT 14
US-10-017-754-966/c
; Sequence 966, Application US/10017754
; Publication No. US20030054363A1
; GENERAL INFORMATION:
;  APPLICANT: Henderson, Robert A.
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Watanabe, Yoshihiro
;  APPLICANT:  Johnson, Jeffrey C.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Marnerakis, Margarita
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  McNabb, Andria
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.478C18
;  CURRENT APPLICATION NUMBER: US/10/017,754
;  CURRENT FILING DATE:  2001-10-29
;  NUMBER OF SEQ ID NOS: 2004
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 966
;   LENGTH: 246
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-017-754-966

  Query Match             6.8%;  Score 246;  DB 14;  Length 246;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-65;
  Matches  246;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2293 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 2352
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy       2353 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 2412
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy       2413 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 2472
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67

Qy       2473 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 2532
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy       2533 GAAAGG 2538
                ||||||
Db          6 GAAAGG 1


RESULT 15
US-10-113-872-966/c
; Sequence 966, Application US/10113872
; Publication No. US20030170255A1
; GENERAL INFORMATION:
;  APPLICANT: Watanabe, Yoshihiro
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.478C19
;  CURRENT APPLICATION NUMBER: US/10/113,872
;  CURRENT FILING DATE:  2002-03-28
;  NUMBER OF SEQ ID NOS: 2011
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 966
;   LENGTH: 246
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-113-872-966

  Query Match             6.8%;  Score 246;  DB 14;  Length 246;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-65;
  Matches  246;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2293 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 2352
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy       2353 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 2412
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy       2413 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 2472
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67

Qy       2473 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 2532
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy       2533 GAAAGG 2538
                ||||||
Db          6 GAAAGG 1


Search completed: March  5, 2004, 13:18:25
Job time : 802.175 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2004, 17:02:19 ; Search time 159.851 Seconds
(without alignments)
12487.583 Million cell updates/sec

Title: US-09-787-097-1
Perfect score: 3597
Sequence: 1 atggtggccgcagcggcggc.....aggtgcaaactgaacaatga 3597

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	3494.4	97.1	4072	3	US-09-245-041-16	Sequence 16, Appl
	2	3476.8	96.7	8589	3	US-09-245-041-14	Sequence 14, Appl
	3	2845.2	79.1	8827	3	US-09-245-041-1	Sequence 1, Appli
	4	2028.8	56.4	2625	3	US-09-245-041-18	Sequence 18, Appl
	5	1428.8	39.7	2419	3	US-09-245-041-8	Sequence 8, Appli
	6	1364.4	37.9	6370	3	US-09-245-041-12	Sequence 12, Appl
	7	463	12.9	1051	3	US-09-245-041-10	Sequence 10, Appl
c	8	246	6.8	246	4	US-09-702-705-966	Sequence 966, App
c	9	246	6.8	246	4	US-09-736-457-966	Sequence 966, App
c	10	246	6.8	246	4	US-09-614-124B-966	Sequence 966, App
c	11	246	6.8	246	4	US-09-671-325-966	Sequence 966, App
	12	212.8	5.9	415	4	US-09-833-381-1916	Sequence 1916, Ap
	13	202.8	5.6	90050	3	US-09-245-041-5	Sequence 5, Appli
c	14	169.2	4.7	17056	3	US-09-245-041-3	Sequence 3, Appli
c	15	132.2	3.7	226	3	US-09-245-041-6	Sequence 6, Appli
c	16	115.4	3.2	221	3	US-09-245-041-7	Sequence 7, Appli
	17	100.2	2.8	5973	3	US-09-245-041-4	Sequence 4, Appli
c	18	48.8	1.4	4403765	3	US-09-103-840A-2	Sequence 2, Appli
c	19	48.8	1.4	4411529	3	US-09-103-840A-1	Sequence 1, Appli
	20	48.6	1.4	832	4	US-09-621-976-2813	Sequence 2813, Ap
	21	46.8	1.3	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c	22	46	1.3	2561	4	US-09-616-289-48	Sequence 48, Appl
	23	45.8	1.3	4403765	3	US-09-103-840A-2	Sequence 2, Appli
	24	44.8	1.2	5329	4	US-09-562-702A-19	Sequence 19, Appl
	25	44.8	1.2	5329	4	US-09-561-818A-19	Sequence 19, Appl
	26	44.8	1.2	5689	4	US-09-562-702A-17	Sequence 17, Appl
	27	44.8	1.2	5689	4	US-09-561-818A-17	Sequence 17, Appl
	28	42.6	1.2	4362	2	US-08-455-073A-1	Sequence 1, Appli
	29	42.4	1.2	700	4	US-09-236-097-7	Sequence 7, Appli
c	30	42	1.2	1213	1	US-08-276-151-6	Sequence 6, Appli
	31	41.6	1.2	538	4	US-09-056-556-180	Sequence 180, App
	32	41.6	1.2	538	4	US-09-072-596-175	Sequence 175, App
	33	41.6	1.2	538	4	US-09-072-967-180	Sequence 180, App
	34	41.4	1.2	1028	3	US-08-118-200-1	Sequence 1, Appli
	35	41.4	1.2	1028	3	US-08-458-745-1	Sequence 1, Appli
	36	41.4	1.2	3789	4	US-09-091-889A-3	Sequence 3, Appli
	37	41.4	1.2	8450	4	US-09-091-889A-5	Sequence 5, Appli
	38	40.6	1.1	1356	4	US-09-252-991A-9233	Sequence 9233, Ap
c	39	40.6	1.1	2715	4	US-09-252-991A-8715	Sequence 8715, Ap
	40	40.6	1.1	3126	4	US-09-252-991A-9066	Sequence 9066, Ap
	41	40.6	1.1	4143	4	US-09-252-991A-9151	Sequence 9151, Ap
c	42	40.6	1.1	4212	4	US-09-252-991A-8929	Sequence 8929, Ap
	43	40.4	1.1	3572	4	US-09-575-574-3	Sequence 3, Appli
	44	40.2	1.1	1026	3	US-07-751-891B-24	Sequence 24, Appl
	45	40.2	1.1	5222	3	US-07-751-891B-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1
US-09-245-041-16
; Sequence 16, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 4072
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-245-041-16

Query Match 97.1%; Score 3494.4; DB 3; Length 4072;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 3541; Conservative 0; Mismatches 41; Indels 2; Gaps 2;

```
Qy      15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74
           || ||| | || | || | | | | | | | |||| || || || ||
Db     234 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAA 293

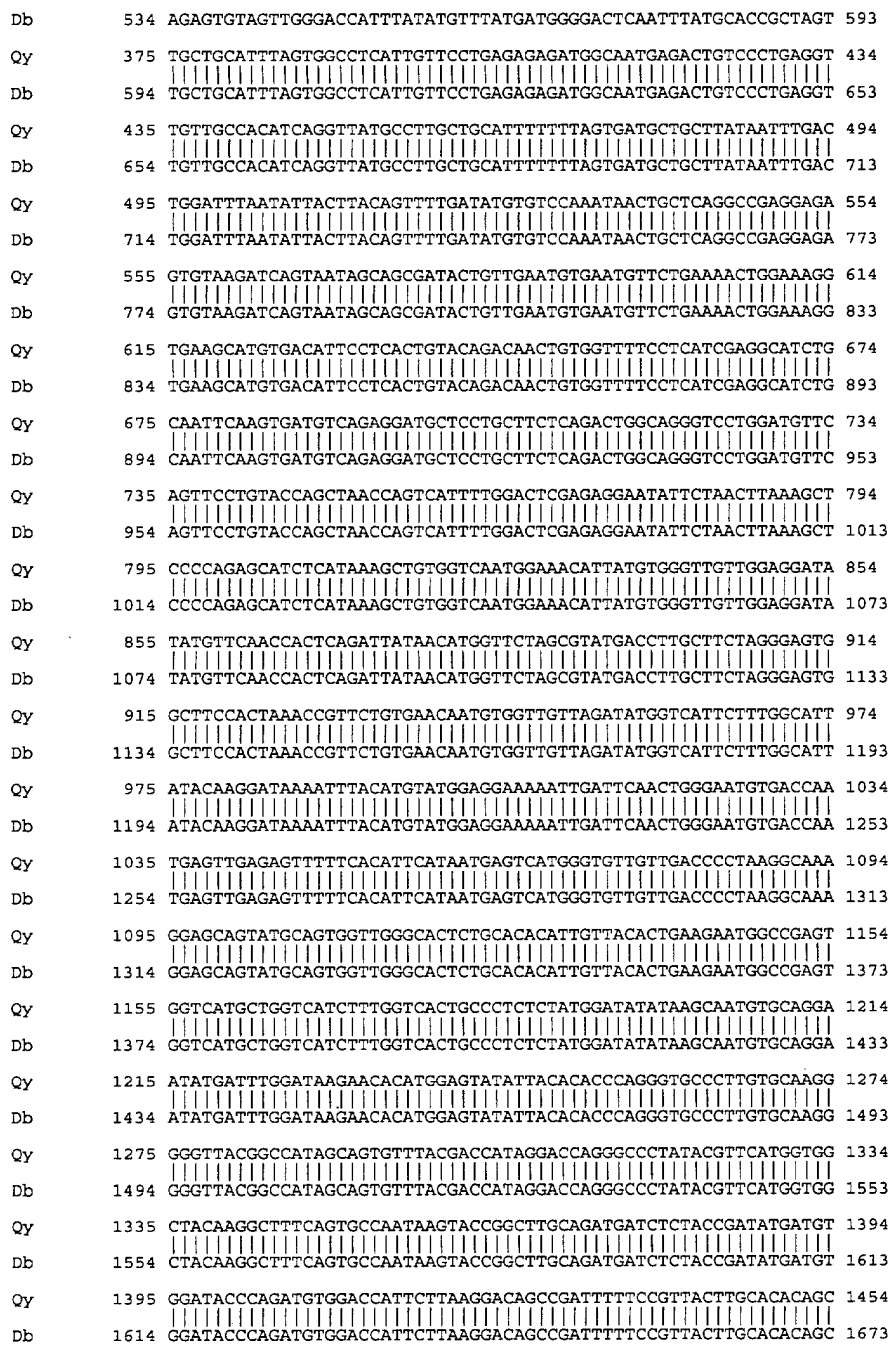
Qy      75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134
           |    | | | || ||||||||||||||||||||||||||||||||||||||||||||
Db     294 GGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 353

Qy     135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     354 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 413

Qy     195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     414 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 473

Qy     255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     474 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 533

Qy     315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```


Db 534 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 593
Qy 375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434
Db 594 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 653
Qy 435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494
Db 654 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 713
Qy 495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554
Db 714 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 773
Qy 555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614
Db 774 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 833
Qy 615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 674
Db 834 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 893
Qy 675 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 734
Db 894 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 953
Qy 735 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 794
Db 954 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 1013
Qy 795 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 854
Db 1014 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 1073
Qy 855 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 914
Db 1074 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 1133
Qy 915 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 974
Db 1134 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 1193
Qy 975 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1034
Db 1194 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1253
Qy 1035 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1094
Db 1254 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1313
Qy 1095 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1154
Db 1314 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1373
Qy 1155 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1214
Db 1374 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1433
Qy 1215 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1274
Db 1434 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1493
Qy 1275 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1334
Db 1494 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1553
Qy 1335 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1394
Db 1554 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1613
Qy 1395 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1454
Db 1614 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1673
Qy 1455 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1514
Db 1674 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1733
Qy 1515 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1574
Db 1734 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1793
Qy 1575 CTGGTCAGTGCTTCCCAGACCTGAT-TCCACCATGATGTCAACAGATTTGGCCATTCCAG 1633
Db 1794 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATT-CAG 1852
Qy 1634 CAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCG 1693
Db 1853 CAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCG 1912
Qy 1694 ACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAG 1753
Db 1913 ACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAG 1972
Qy 1754 CAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGG 1813
Db 1973 CAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGG 2032
Qy 1814 CGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTC 1873
Db 2033 CGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTC 2092
Qy 1874 TTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATG 1933
Db 2093 TTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATG 2152
Qy 1934 ACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGA 1993
Db 2153 ACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGA 2212
Qy 1994 TCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGA 2053
Db 2213 TCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGA 2272
Qy 2054 AGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGG 2113
Db 2273 AGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGG 2332
Qy 2114 AGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCAT 2173
Db 2333 AGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCAT 2392
Qy 2174 GTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACC 2233
Db 2393 GTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACC 2452
Qy 2234 ACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGC 2293
Db 2453 ACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGC 2512
Qy 2294 TGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCC 2353
Db 2513 TGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCC 2572
Qy 2354 TTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTT 2413
Db 2573 TTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTT 2632
Qy 2414 TACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAAC 2473
Db 2633 TACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAAC 2692
Qy 2474 CCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTG 2533
Db 2693 CCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTG 2752

```
Qy       2534 AAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCAT 2593
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2753 AAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCAT 2812

Qy       2594 GTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTG 2653
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2813 GTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTG 2872

Qy       2654 TGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGA 2713
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2873 TGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGA 2932

Qy       2714 TGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGC 2773
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2933 TGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGC 2992

Qy       2774 AACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGG 2833
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2993 AACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGG 3052

Qy       2834 GTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCAC 2893
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3053 GTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCAC 3112

Qy       2894 AGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTC 2953
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3113 AGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTC 3172

Qy       2954 ACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGA 3013
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3173 ACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGA 3232

Qy       3014 AGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTG 3073
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3233 AGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTG 3292

Qy       3074 ATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCA 3133
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3293 ATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCA 3352

Qy       3134 ACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGC 3193
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3353 ACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGC 3412

Qy       3194 TATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTC 3253
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3413 TATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTC 3472

Qy       3254 TTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAG 3313
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3473 TTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAG 3532

Qy       3314 CTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATG 3373
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3533 CTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATG 3592

Qy       3374 CCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGG 3433
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3593 CCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGG 3652

Qy       3434 CTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCT 3493
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3653 CTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCT 3712

Qy       3494 CTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTA 3553
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3713 CTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTA 3772

Qy       3554 ATTTCACCTGGCCCATCAAAATTCAGGTGCAAACTGAACAATGA 3597
             |||||||||||||||||||||||||||||||||||||||||||
Db       3773 ATTTCACCTGGCCCATCAAAATTCAGGTGCAAACTGAACAATGA 3816
```

```
RESULT 2
US-09-245-041-14
; Sequence 14, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT: Nagle, D.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
;  TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
;  FILE REFERENCE: 7853-136
;  CURRENT APPLICATION NUMBER: US/09/245,041
;  CURRENT FILING DATE: 1999-02-05
;  EARLIER APPLICATION NUMBER: 60/093,630
;  EARLIER FILING DATE: 1998-07-21
;  EARLIER APPLICATION NUMBER: 60/104,978
;  EARLIER FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 131
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 14
;   LENGTH: 8589
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-245-041-14

  Query Match            96.7%;  Score 3476.8;  DB 3;  Length 8589;
  Best Local Similarity  98.8%;  Pred. No. 0;
  Matches 3524;  Conservative    0;  Mismatches   42;  Indels    2;  Gaps    2;

Qy         15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74
              || |||    |  ||   | ||  |  |  |  | | |   |||| ||   || || ||
Db        234 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAA 293

Qy         75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134
              |  | | |  ||  ||||||||||||||||||||||||||||||||||||||||||||||
Db        294 GGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 353

Qy        135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        354 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 413

Qy        195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 473

Qy        255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 533

Qy        315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 593

Qy        375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 653

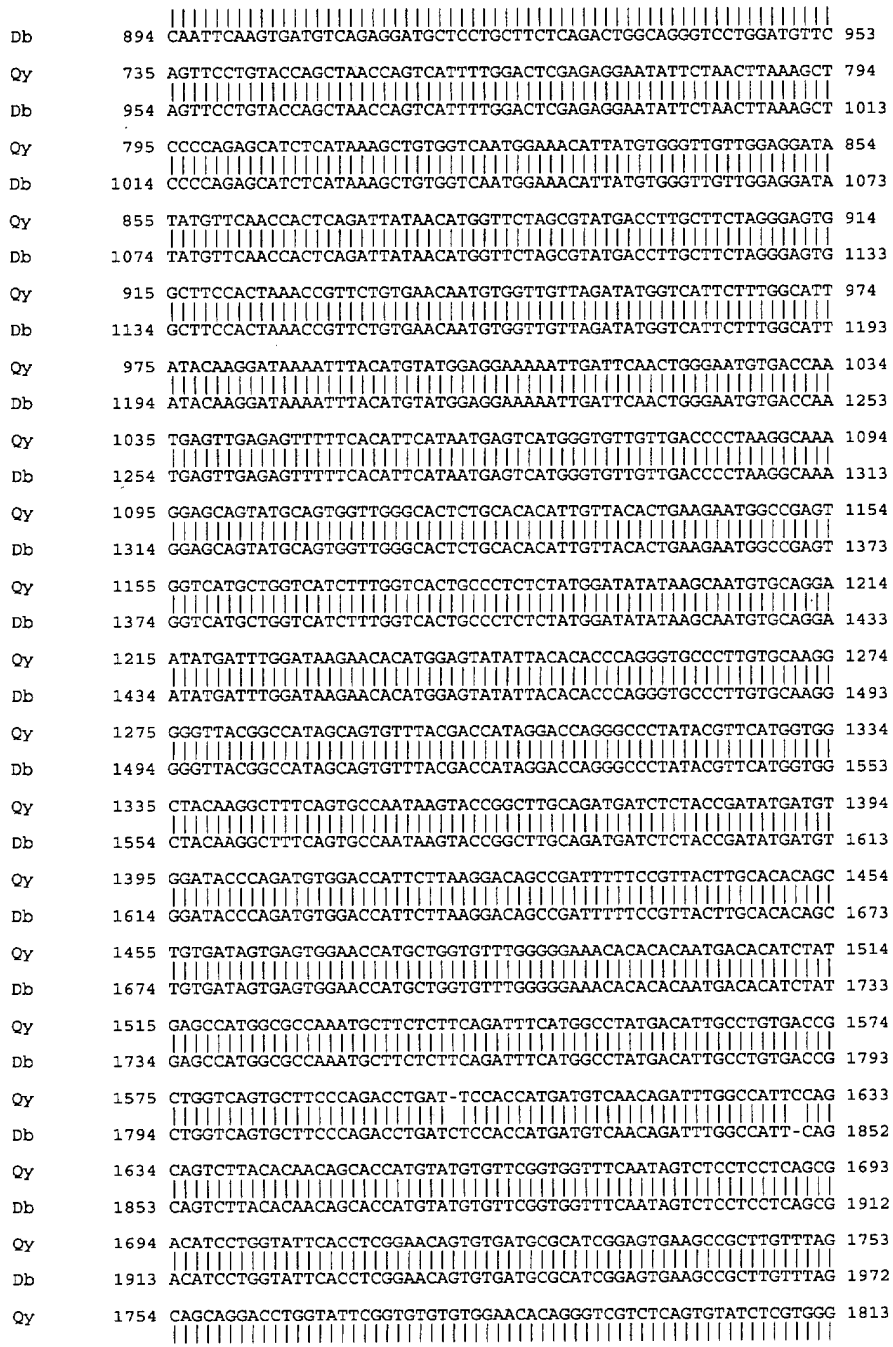
Qy        435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        654 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 713

Qy        495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        714 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 773

Qy        555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        774 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 833

Qy        615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        834 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 893

Qy        675 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 734
```


Db 894 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 953
Qy 735 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 794
Db 954 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 1013
Qy 795 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 854
Db 1014 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 1073
Qy 855 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 914
Db 1074 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 1133
Qy 915 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 974
Db 1134 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 1193
Qy 975 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1034
Db 1194 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1253
Qy 1035 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1094
Db 1254 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1313
Qy 1095 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1154
Db 1314 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1373
Qy 1155 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1214
Db 1374 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1433
Qy 1215 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1274
Db 1434 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1493
Qy 1275 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1334
Db 1494 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1553
Qy 1335 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1394
Db 1554 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1613
Qy 1395 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1454
Db 1614 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1673
Qy 1455 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1514
Db 1674 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1733
Qy 1515 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1574
Db 1734 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1793
Qy 1575 CTGGTCAGTGCTTCCCAGACCTGAT-TCCACCATGATGTCAACAGATTTGGCCATTCCAG 1633
Db 1794 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATT-CAG 1852
Qy 1634 CAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCG 1693
Db 1853 CAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCG 1912
Qy 1694 ACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAG 1753
Db 1913 ACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAG 1972
Qy 1754 CAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGG 1813
Db 1973 CAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGG 2032
Qy 1814 CGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTC 1873
Db 2033 CGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTC 2092
Qy 1874 TTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATG 1933
Db 2093 TTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATG 2152
Qy 1934 ACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGA 1993
Db 2153 ACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGA 2212
Qy 1994 TCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGA 2053
Db 2213 TCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGA 2272
Qy 2054 AGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGG 2113
Db 2273 AGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGG 2332
Qy 2114 AGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCAT 2173
Db 2333 AGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCAT 2392
Qy 2174 GTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACC 2233
Db 2393 GTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACC 2452
Qy 2234 ACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGC 2293
Db 2453 ACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGC 2512
Qy 2294 TGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCC 2353
Db 2513 TGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCC 2572
Qy 2354 TTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTT 2413
Db 2573 TTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTT 2632
Qy 2414 TACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAAC 2473
Db 2633 TACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAAC 2692
Qy 2474 CCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTG 2533
Db 2693 CCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTG 2752
Qy 2534 AAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCAT 2593
Db 2753 AAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCAT 2812
Qy 2594 GTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTG 2653
Db 2813 GTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTG 2872
Qy 2654 TGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGA 2713
Db 2873 TGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGA 2932
Qy 2714 TGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGC 2773
Db 2933 TGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGC 2992
Qy 2774 AACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGG 2833
Db 2993 AACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGG 3052
Qy 2834 GTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCAC 2893
Db 3053 GTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCAC 3112

```
Qy       2894 AGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTC 2953
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3113 AGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTC 3172

Qy       2954 ACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGA 3013
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3173 ACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGA 3232

Qy       3014 AGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTG 3073
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3233 AGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTG 3292

Qy       3074 ATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCA 3133
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3293 ATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCA 3352

Qy       3134 ACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGC 3193
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3353 ACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGC 3412

Qy       3194 TATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTC 3253
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3413 TATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTC 3472

Qy       3254 TTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAG 3313
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3473 TTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAG 3532

Qy       3314 CTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATG 3373
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3533 CTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATG 3592

Qy       3374 CCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGG 3433
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3593 CCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGG 3652

Qy       3434 CTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCT 3493
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3653 CTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCT 3712

Qy       3494 CTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTA 3553
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3713 CTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTA 3772

Qy       3554 ATTTCACCTGGCCCATCAAAATTCAGGT 3581
             ||||||||||||||||||||||||||| |
Db       3773 ATTTCACCTGGCCCATCAAAATTCAGAT 3800
```

```
RESULT 3
US-09-245-041-1
; Sequence 1, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT: Nagle, D.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
;  TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
;  FILE REFERENCE: 7853-136
;  CURRENT APPLICATION NUMBER: US/09/245,041
;  CURRENT FILING DATE: 1999-02-05
;  EARLIER APPLICATION NUMBER: 60/093,630
;  EARLIER FILING DATE: 1998-07-21
;  EARLIER APPLICATION NUMBER: 60/104,978
;  EARLIER FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 131
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 8827
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-245-041-1
```

```
Query Match           79.1%;  Score 2845.2;  DB 3;  Length 8827;
Best Local Similarity 90.3%;  Pred. No. 0;
Matches 3064;  Conservative    0;  Mismatches  328;  Indels    2;  Gaps    2;

Qy        189 ACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAA 248
              |||||||| |||||||| ||||| |||||||||||||| |||||||||| || |||||
Db        677 ACTAACTGGCTCTTCTGGATTTGTAACAGATGGACCTGGGAATTATAAATATAAGACGAA 736

Qy        249 GTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTT 308
              |||||| |||||||||||||||||||||||||||||||||||||| ||||| |||||
Db        737 GTGCACATGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGCTTCAACCATTT 796

Qy        309 TGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACC 368
              ||||||||| ||||| |||||||||||||||||||||||||||||||||||| || |||||
Db        797 TGCTACAGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATCTACGCACC 856

Qy        369 GCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCC 428
              ||  |||||| |||||||||||||||||||||| |||||||||||||||||| |  ||
Db        857 TCTGATTGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCTCC 916

Qy        429 TGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAA 488
              ||||||  || ||| |||||||||||| |||||||||||||| |||||||||||||||||
Db        917 TGAGGTCACTGTCACTTCAGGTTATGCACTGCTGCATTTTTTCAGTGATGCTGCTTATAA 976

Qy        489 TTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCG 548
              | |||||||||||||||| |||||| |||||| |||||||| ||||| ||||||||||
Db        977 TCTGACTGGATTTAATATCACTTACAATTTTGACATGTGTCCGAATAATTGCTCAGGCCG 1036

Qy        549 AGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTG 608
              |||||||||||||| |||||| ||||||     |||||| |||||||||||||||||||
Db       1037 AGGAGAGTGTAAGAGCAGTAACAGCAGCAGCGCTGTTGAGTGTGAATGTTCTGAAAACTG 1096

Qy        609 GAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGG 668
              |||||| ||  | |||||||||||||||||||||||||||||||||| |||||||| ||||
Db       1097 GAAAGGGGAGTCGTGTGACATTCCTCACTGTACAGACAACTGTGGCTTTCCTCACCGAGG 1156

Qy        669 CATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGG 728
              |||||| ||| |||| ||| |||||| |||||||||| |  ||||||||||||||||||
Db       1157 CATCTGTAATGCAAGCGATACCAGAGGGTGCTCCTGCTTTCCTCACTGGCAGGGTCCTGG 1216

Qy        729 ATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTT 788
              ||||||| |||||| |||||||||||||| |||||||||||||| |||||||| | ||
Db       1217 ATGTTCAATTCCTGTGCCAGCTAACCAGTCTTTTTGGACTCGAGAAGAATATTCTGATTT 1276

Qy        789 AAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGG 848
              |||||| |||||||||||||||||||||||||||||||||||||||| || |||||||||
Db       1277 AAAGCTTCCCAGAGCCTCTCATAAAGCTGTGGTCAATGGAAATATAATGTGGGTTGTTGG 1336

Qy        849 AGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAG 908
              |||||||||||||||| |||||||| | ||||||| |||||||||||||  |||||||
Db       1337 CGGATATATGTTCAACCATTCAGATTACAGCATGGTTTTAGCGTATGACCTGACTTCTAG 1396

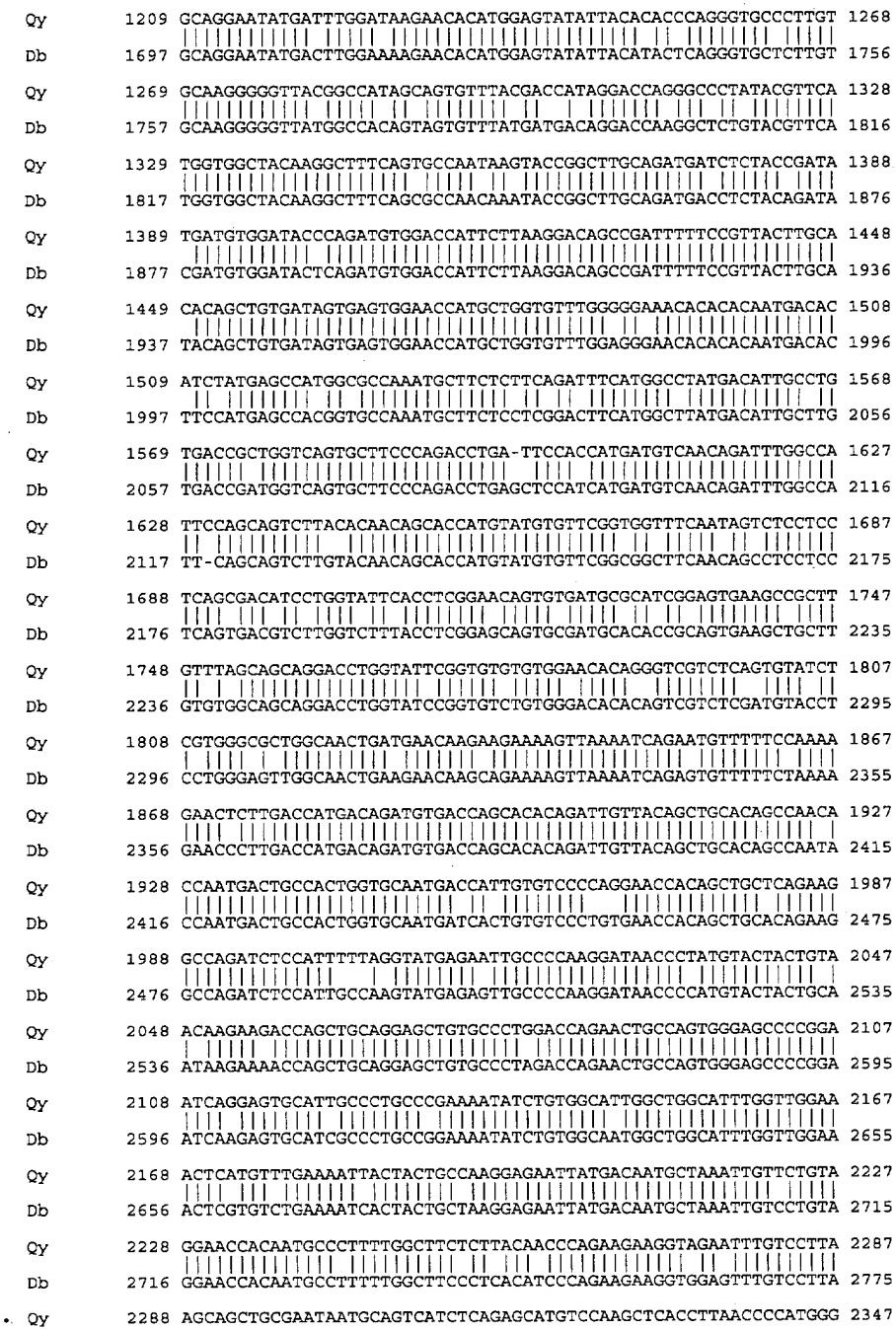
Qy        909 GGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTT 968
              ||| |||||||||||||| |||||||||| |||||||| |||||||||||||||||||||
Db       1397 GGAATGGCTTCCACTAAACCATTCTGTGAACAGTGTGGTTGTAAGATATGGTCATTCTTT 1456

Qy        969 GGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGT 1028
              ||||||| | |||||||||| |||||||||||||||||||||||||||| ||||| ||
Db       1457 GGCATTACATAAGGATAAAATCTACATGTATGGAGGAAAAATTGATTCAACAGGGAACGT 1516

Qy       1029 GACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAA 1088
              |||||||||| ||||||| ||||| |||||||||| |||||||| ||||| || || ||
Db       1517 GACCAATGAGCTGAGAGTATTTCATATTCATAATGAATCATGGGTATTGTTAACTCCGAA 1576

Qy       1089 GGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGG 1148
              || ||||| |||||||||||||||||| ||||| |||||||||||||||||     |||
Db       1577 AGCTAAGGATCAGTATGCAGTGGTTGGACACTCAGCACACATTGTTACACTGGCATCTGG 1636

Qy       1149 CCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGT 1208
              ||| |||||||||| |||||||||| ||||| ||||| |||||||||||||||||  |||
Db       1637 CCGTGTGGTCATGTTGGTCATCTTCGGTCATTGCCCACTCTATGGATATATAAGCGTTGT 1696
```


Qy 1209 GCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGT 1268
Db 1697 GCAGGAATATGACTTGGAAAAGAACACATGGAGTATATTACATACTCAGGGTGCTCTTGT 1756
Qy 1269 GCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCA 1328
Db 1757 GCAAGGGGGTTATGGCCACAGTAGTGTTTATGATGACAGGACCAAGGCTCTGTACGTTCA 1816
Qy 1329 TGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATA 1388
Db 1817 TGGTGGCTACAAGGCTTTCAGCGCCAACAAATACCGGCTTGCAGATGACCTCTACAGATA 1876
Qy 1389 TGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCA 1448
Db 1877 CGATGTGGATACTCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCA 1936
Qy 1449 CACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACAC 1508
Db 1937 TACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGAGGGAACACACACAATGACAC 1996
Qy 1509 ATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTG 1568
Db 1997 TTCCATGAGCCACGGTGCCAAATGCTTCTCCTCGGACTTCATGGCTTATGACATTGCTTG 2056
Qy 1569 TGACCGCTGGTCAGTGCTTCCCAGACCTGA-TTCCACCATGATGTCAACAGATTTGGCCA 1627
Db 2057 TGACCGATGGTCAGTGCTTCCCAGACCTGAGCTCCATCATGATGTCAACAGATTTGGCCA 2116
Qy 1628 TTCCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCC 1687
Db 2117 TT-CAGCAGTCTTGTACAACAGCACCATGTATGTGTTCGGCGGCTTCAACAGCCTCCTCC 2175
Qy 1688 TCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTT 1747
Db 2176 TCAGTGACGTCTTGGTCTTTACCTCGGAGCAGTGCGATGCACACCGCAGTGAAGCTGCTT 2235
Qy 1748 GTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCT 1807
Db 2236 GTGTGGCAGCAGGACCTGGTATCCGGTGTCTGTGGGACACACAGTCGTCTCGATGTACCT 2295
Qy 1808 CGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAA 1867
Db 2296 CCTGGGAGTTGGCAACTGAAGAACAAGCAGAAAAGTTAAAATCAGAGTGTTTTTCTAAAA 2355
Qy 1868 GAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACA 1927
Db 2356 GAACCCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAATA 2415
Qy 1928 CCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAG 1987
Db 2416 CCAATGACTGCCACTGGTGCAATGATCACTGTGTCCCTGTGAACCACAGCTGCACAGAAG 2475
Qy 1988 GCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTA 2047
Db 2476 GCCAGATCTCCATTGCCAAGTATGAGAGTTGCCCCAAGGATAACCCCATGTACTACTGCA 2535
Qy 2048 ACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGA 2107
Db 2536 ATAAGAAAACCAGCTGCAGGAGCTGTGCCCTAGACCAGAACTGCCAGTGGGAGCCCCGGA 2595
Qy 2108 ATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAA 2167
Db 2596 ATCAAGAGTGCATCGCCCTGCCGGAAAATATCTGTGGCAATGGCTGGCATTTGGTTGGAA 2655
Qy 2168 ACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTA 2227
Db 2656 ACTCGTGTCTGAAAATCACTACTGCTAAGGAGAATTATGACAATGCTAAATTGTCCTGTA 2715
Qy 2228 GGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTA 2287
Db 2716 GGAACCACAATGCCTTTTTGGCTTCCCTCACATCCCAGAAGAAGGTGGAGTTTGTCCTTA 2775
Qy 2288 AGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGG 2347
Db 2776 AGCAGCTTCGATTAATGCAATCATCTCAAAGTATGTCCAAGCTCACTCTGACTCCATGGG 2835
Qy 2348 TCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAA 2407
Db 2836 TTGGTCTTCGGAAGATCAATGTGTCTTACTGGTGCTGGGAGGATATGTCTCCATTCACAA 2895
Qy 2408 ATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTAT 2467
Db 2896 ATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGTGATGCTGGCTTCTGTGGGATCTTGT 2955
Qy 2468 CAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTG 2527
Db 2956 CAGAGCCTAGTACTCGGGGATTAAAGGCTGCAACCTGCATCAACCCTCTCAATGGCAGCG 3015
Qy 2528 TCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGA 2587
Db 3016 TCTGTGAAAGGCCTGCAAACCACAGTGCCAAGCAGTGCCGGACACCATGTGCCCTGCGGA 3075
Qy 2588 CAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGC 2647
Db 3076 CAGCGTGTGGCGAGTGCACTAGCAGCAGCTCGGAGTGCATGTGGTGCAGTAACATGAAGC 3135
Qy 2648 AGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGT 2707
Db 3136 AGTGTGTGGACTCCAATGCCTACGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGT 3195
Qy 2708 ATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCT 2767
Db 3196 ATACGATGAGCAGCTGCCCACCTGAAAATTGCTCTGGCTACTGTACCTGCAGCCATTGCT 3255
Qy 2768 TGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCA 2827
Db 3256 TGGAGCAGCCAGGCTGTGGTTGGTGTACTGATCCTAGCAATACTGGGAAAGGAAAATGTA 3315
Qy 2828 TAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCT 2887
Db 3316 TTGAGGGCAGCTATAAAGGACCTGTGAAGATGCCGTCACAGGCCTCTGCAGGAAATGTGT 3375
Qy 2888 ATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTT 2947
Db 3376 ATCCACAGCCCCTTCTGAACTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTT 3435
Qy 2948 TCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCT 3007
Db 3436 TCATTCACTGTCCAGCTTGCCAGTGCAACGGACACAGCAAATGCATCAACCAGAGTATCT 3495
Qy 3008 GTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCT 3067
Db 3496 GTGAGAAGTGTGAGGACCTGACCACGGGCAAGCACTGCGAGACCTGCATATCTGGCTTCT 3555
Qy 3068 ACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTC 3127
Db 3556 ATGGTGACCCGACTAATGGAGGCAAATGTCAGCCATGCAAGTGCAATGGGCACGCATCAC 3615
Qy 3128 TGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGT 3187
Db 3616 TGTGCAACACCAACACCGGCAAGTGCTTCTGTACCACCAAAGGTGTCAAGGGGGACGAGT 3675
Qy 3188 GCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATT 3247
Db 3676 GCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAAAGGAACATGCTACT 3735
Qy 3248 ATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATT 3307
Db 3736 ATACCCTTCTCATTGACTATCAGTTCACCTTTAGCCTGTCCCAGGAAGACGACCGCTACT 3795
Qy 3308 ACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCA 3367
Db 3796 ACACAGCCATCAACTTTGTGGCTACTCCTGATGAACAAAACAGGGATTTGGACATGTTCA 3855
Qy 3368 TCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAA 3427

```
Db       3856 TCAATGCCTCCAAAAACTTCAACCTCAACATCACCTGGGCCACCAGCTTCCCAGCCGGAA 3915

Qy       3428 CCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATA 3487

Db       3916 CCCAGACTGGAGAAGAGGTGCCTGTTGTTTCAAAAACCAACATCAAGGAATACAAAGATA 3975

Qy       3488 GTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATG 3547

Db       3976 GCTTCTCTAATGAGAAATTTGATTTTCGCAACCATCCAAACATCACTTTCTTTGTTTATG 4035

Qy       3548 TCAGTAATTTCACCTGGCCCATCAAAATTCAGGT 3581

Db       4036 TCAGTAATTTCACTTGGCCCATCAAAATTCAGAT 4069


RESULT 4
US-09-245-041-18
; Sequence 18, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT: Nagle, D.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
;  TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
;  FILE REFERENCE: 7853-136
;  CURRENT APPLICATION NUMBER: US/09/245,041
;  CURRENT FILING DATE: 1999-02-05
;  EARLIER APPLICATION NUMBER: 60/093,630
;  EARLIER FILING DATE: 1998-07-21
;  EARLIER APPLICATION NUMBER: 60/104,978
;  EARLIER FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 131
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 18
;   LENGTH: 2625
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-245-041-18

  Query Match             56.4%;  Score 2028.8;  DB 3;  Length 2625;
  Best Local Similarity   97.0%;  Pred. No. 0;
  Matches 2088;  Conservative    0;  Mismatches   62;  Indels    2;  Gaps    2;

Qy         15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74

Db        234 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAA 293

Qy         75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134

Db        294 GGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 353

Qy        135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194

Db        354 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 413

Qy        195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254

Db        414 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 473

Qy        255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314

Db        474 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 533

Qy        315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374

Db        534 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 593

Qy        375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434

Db        594 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 653

Qy        435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494

Db        654 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 713

Qy        495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554

Db        714 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 773

Qy        555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614

Db        774 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 833

Qy        615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 674

Db        834 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 893

Qy        675 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 734

Db        894 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 953

Qy        735 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 794

Db        954 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 1013

Qy        795 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 854

Db       1014 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 1073

Qy        855 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 914

Db       1074 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 1133

Qy        915 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 974

Db       1134 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 1193

Qy        975 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1034

Db       1194 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1253

Qy       1035 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1094

Db       1254 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1313

Qy       1095 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1154

Db       1314 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1373

Qy       1155 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1214

Db       1374 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1433

Qy       1215 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1274

Db       1434 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1493

Qy       1275 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1334

Db       1494 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1553

Qy       1335 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1394

Db       1554 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1613

Qy       1395 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1454

Db       1614 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1673

Qy       1455 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1514

Db       1674 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1733

Qy       1515 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1574

Db       1734 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1793
```

```
Qy       1575 CTGGTCAGTGCTTCCCAGACCTGAT-TCCACCATGATGTCAACAGATTTGGCCATTCCAG 1633
             |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||| |||
Db       1794 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATT-CAG 1852

Qy       1634 CAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCG 1693
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1853 CAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCG 1912

Qy       1694 ACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAG 1753
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1913 ACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAG 1972

Qy       1754 CAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGG 1813
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1973 CAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGG 2032

Qy       1814 CGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTC 1873
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2033 CGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTC 2092

Qy       1874 TTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATG 1933
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2093 TTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATG 2152

Qy       1934 ACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGA 1993
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2153 ACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGA 2212

Qy       1994 TCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGA 2053
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2213 TCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGA 2272

Qy       2054 AGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGG 2113
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2273 AGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGG 2332

Qy       2114 AGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGG 2165
             |||||||||||||||||  |   ||  |  |  || |  |  || ||
Db       2333 AGTGCATTGCCCTGCCCGGTAGGCCTTGCAGGGTCATCTTGGTGTGTGTGGG 2384
```

RESULT 5
US-09-245-041-8
; Sequence 8, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 2419
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-8

Query Match 39.7%; Score 1428.8; DB 3; Length 2419;
Best Local Similarity 89.4%; Pred. No. 0;
Matches 1560; Conservative 0; Mismatches 182; Indels 2; Gaps 2;

```
Qy        189 ACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAA 248
              ||||||||| |||||||| ||||| |||||||||||||| |||||||||| || ||||||
Db        677 ACTAACTGGCTCTTCTGGATTTGTAACAGATGGACCTGGGAATTATAAATATAAGACGAA 736

Qy        249 GTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTT 308
              |||||| |||||||||||||||||||||||||||||||||||||||||||| ||||| |||||
Db        737 GTGCACATGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGCTTCAACCATTT 796

Qy        309 TGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACC 368
              |||||||| ||||| |||||||||||||||||||||||||||||||||||| || |||||
Db        797 TGCTACAGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATCTACGCACC 856

Qy        369 GCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCC 428
              ||  ||||||| |||||||||||||||||||| |||||||||||||||| | ||
Db        857 TCTGATTGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCTCC 916

Qy        429 TGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAA 488
              ||||||   || ||| |||||||||||  |||||||||||| ||||||||||||||||||
Db        917 TGAGGTCACTGTCACTTCAGGTTATGCACTGCTGCATTTTTTCAGTGATGCTGCTTATAA 976

Qy        489 TTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCG 548
              | |||||||||||||||| |||||| |||||| |||||| |||||||||||||||||||
Db        977 TCTGACTGGATTTAATATCACTTACAATTTTGACATGTGTCCGAATAATTGCTCAGGCCG 1036

Qy        549 AGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTG 608
              |||||||||||||| |||||| ||||||     |||||| ||||||||||||||||||||
Db       1037 AGGAGAGTGTAAGAGCAGTAACAGCAGCAGCGCTGTTGAGTGTGAATGTTCTGAAAACTG 1096

Qy        609 GAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGG 668
              |||||  ||   | |||||||||||||||||||||||||||||||| |||||| ||||
Db       1097 GAAAGGGGAGTCGTGTGACATTCCTCACTGTACAGACAACTGTGGCTTTCCTCACCGAGG 1156

Qy        669 CATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGG 728
              |||||| ||| |||| ||| |||||| |||||||||| |  ||||||||||||||||
Db       1157 CATCTGTAATGCAAGCGATACCAGAGGGTGCTCCTGCTTTCCTCACTGGCAGGGTCCTGG 1216

Qy        729 ATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTT 788
              ||||||| ||||||| |||||||||||||| |||||||||||||| |||||||| | ||
Db       1217 ATGTTCAATTCCTGTGCCAGCTAACCAGTCTTTTTGGACTCGAGAAGAATATTCTGATTT 1276

Qy        789 AAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGG 848
              |||||| |||||||||||||||||||||||||||||||||||||||| || ||||||||||
Db       1277 AAAGCTTCCCAGAGCCTCTCATAAAGCTGTGGTCAATGGAAATATAATGTGGGTTGTTGG 1336

Qy        849 AGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAG 908
              |||||||||||||||| |||||||| | ||||||| |||||||||||||  |||||||
Db       1337 CGGATATATGTTCAACCATTCAGATTACAGCATGGTTTTAGCGTATGACCTGACTTCTAG 1396

Qy        909 GGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTT 968
              ||| |||||||||||||| |||||||||| |||||||| ||||||||||||||||||||
Db       1397 GGAATGGCTTCCACTAAACCATTCTGTGAACAGTGTGGTTGTAAGATATGGTCATTCTTT 1456

Qy        969 GGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGT 1028
              ||||||| | |||||||||||| |||||||||||||||||||||||||||||| ||||| ||
Db       1457 GGCATTACATAAGGATAAAATCTACATGTATGGAGGAAAAATTGATTCAACAGGGAACGT 1516

Qy        1029 GACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAA 1088
               |||||||||| ||||||| ||||| |||||||||| |||||||| ||||| || || ||
Db        1517 GACCAATGAGCTGAGAGTATTTCATATTCATAATGAATCATGGGTATTGTTAACTCCGAA 1576

Qy        1089 GGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGG 1148
               || ||||| |||||||||||||||| |||||| |||||||||||||||||      |||
Db        1577 AGCTAAGGATCAGTATGCAGTGGTTGGACACTCAGCACACATTGTTACACTGGCATCTGG 1636

Qy        1149 CCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGT 1208
               ||| |||||||| |||||||||| |||||  ||||| |||||||||||||||||  |||
Db        1637 CCGTGTGGTCATGTTGGTCATCTTCGGTCATTGCCCACTCTATGGATATATAAGCGTTGT 1696

Qy        1209 GCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGT 1268
               |||||||||||| ||||| |||||||||||||||||||||| || |||||||| |||||
Db        1697 GCAGGAATATGACTTGGAAAAGAACACATGGAGTATATTACATACTCAGGGTGCTCTTGT 1756

Qy        1269 GCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCA 1328
               |||||||||||| ||||| || |||||||| ||  | ||||||| ||| || ||||||||
Db        1757 GCAAGGGGGTTATGGCCACAGTAGTGTTTATGATGACAGGACCAAGGCTCTGTACGTTCA 1816
```

```
Qy       1329 TGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATA 1388
             ||||||||||||||||||||| ||||| || |||||||||||||||| |||||| ||||
Db       1817 TGGTGGCTACAAGGCTTTCAGCGCCAACAAATACCGGCTTGCAGATGACCTCTACAGATA 1876

Qy       1389 TGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCA 1448
             |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1877 CGATGTGGATACTCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCA 1936

Qy       1449 CACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACAC 1508
             |||||||||||||||||||||||||||||||||||||| || ||||||||||||||||||
Db       1937 TACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGAGGGAACACACACAATGACAC 1996

Qy       1509 ATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTG 1568
             || |||||||| || |||||||||||||| || || |||||||| |||||||||||| ||
Db       1997 TTCCATGAGCCACGGTGCCAAATGCTTCTCCTCGGACTTCATGGCTTATGACATTGCTTG 2056

Qy       1569 TGACCGCTGGTCAGTGCTTCCCAGACCTGA-TTCCACCATGATGTCAACAGATTTGGCCA 1627
             |||||| |||||||||||||||||||||||  |||| |||||||||||||||||||||||
Db       2057 TGACCGATGGTCAGTGCTTCCCAGACCTGAGCTCCATCATGATGTCAACAGATTTGGCCA 2116

Qy       1628 TTCCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCC 1687
             || |||||||||| |||||||||||||||||||||||| || ||||| || |||||||
Db       2117 TT-CAGCAGTCTTGTACAACAGCACCATGTATGTGTTCGGCGGCTTCAACAGCCTCCTCC 2175

Qy       1688 TCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTT 1747
             |||| ||| || |||| || |||||||| ||||| ||||| || || |||||||| ||||
Db       2176 TCAGTGACGTCTTGGTCTTTACCTCGGAGCAGTGCGATGCACACCGCAGTGAAGCTGCTT 2235

Qy       1748 GTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCT 1807
             || | |||||||||||||||||| |||||| |||| |||||  |||||||  |||| ||
Db       2236 GTGTGGCAGCAGGACCTGGTATCCGGTGTCTGTGGGACACACAGTCGTCTCGATGTACCT 2295

Qy       1808 CGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAA 1867
             | |||| | |||||||||| |||||||  |||||||||||||||||||  |||||||| ||||
Db       2296 CCTGGGAGTTGGCAACTGAAGAACAAGCAGAAAAGTTAAAATCAGAGTGTTTTTCTAAAA 2355

Qy       1868 GAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACA 1927
             |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       2356 GAACCCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAATA 2415

Qy       1928 CCAA 1931
             ||||
Db       2416 CCAA 2419
```

```
RESULT 6
US-09-245-041-12
; Sequence 12, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT: Nagle, D.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
;  TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
;  FILE REFERENCE: 7853-136
;  CURRENT APPLICATION NUMBER: US/09/245,041
;  CURRENT FILING DATE: 1999-02-05
;  EARLIER APPLICATION NUMBER: 60/093,630
;  EARLIER FILING DATE: 1998-07-21
;  EARLIER APPLICATION NUMBER: 60/104,978
;  EARLIER FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 131
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 12
;   LENGTH: 6370
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-245-041-12

  Query Match             37.9%;  Score 1364.4;  DB 3;  Length 6370;
  Best Local Similarity   95.0%;  Pred. No. 0;
  Matches 1492;  Conservative    7;  Mismatches   48;  Indels   23;  Gaps    8;
```

```
Qy       2035 ATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAG 2094
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAG 60

Qy       2095 TGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGG 2154
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGG 120

Qy       2155 CATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCT 2214
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 CATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCT 180

Qy       2215 AAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTA 2274
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 AAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTA 240

Qy       2275 GAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACC 2334
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 GAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACC 300

Qy       2335 TTAACCCCATGGGTCGGCCTT---CGGAAGATCAATGTGTCCTACTGG-----TGCTGGG 2386
             |||||||||||||||| |  | | | :|||||||||||||||:|     |  |  |
Db         301 TTAACCCCATGGGTCGGGCCTTCGGGAAGGTYCAATGTGTCCTACTKGGTGCTGGGGAAG 360

Qy       2387 AAGATATGTCCCCATTTACAAATAG-TTTACTACAGTGGATGCCGTCTG----AGCCCAG 2441
             | || :  |||  ||||||||||| |||||||||||||     :   |      ||||||
Db         361 GATATGKTCCCCATTTTACAAATAGTTTTACTACAGTGGGATGSCCGTCTTGAGGCCCAG 420

Qy       2442 TGATGCT--GGATTCTGTGGAATTTT-----ATCAGAACCCAGTACTCGGGGACTGAAGG 2494
             || ||||  |:|||||||| |  |||     |  |  |       ||||||||||||
Db         421 TGTTGCTTGGRATTCTGTGGGAATTTTATTCAGGAACCCAGTTACTTCGGGGACTGAAGG 480

Qy       2495 CTGCAACCTGCAT--CAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAG 2552
             |||||||||||||  | ||  ::|||||||||||||||||||||||||||||||||||||
Db         481 CTGCAACCTGCATTCAACCCACTYMAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAG 540

Qy       2553 TGCTAA-GCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCG 2611
             ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 TGCTAAGGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCG 600

Qy       2612 GCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATG 2671
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 GCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATG 660

Qy       2672 TGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTG 2731
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 TGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTG 720

Qy       2732 AAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGT 2791
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 AAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGT 780

Qy       2792 GTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAG 2851
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 GTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAG 840

Qy       2852 TGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCA 2911
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         841 TGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCA 900

Qy       2912 GCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAAT 2971
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 GCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAAT 960

Qy       2972 GCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCA 3031
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 GCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCA 1020

Qy       3032 CAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGA 3091
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1021 CAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGA 1080

Qy       3092 AATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGT 3151
```

```
Db       1081 AATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGT 1140

Qy       3152 GCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATC 3211

Db       1141 GCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATC 1200

Qy       3212 GATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGT 3271

Db       1201 GATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGT 1260

Qy       3272 TCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTA 3331

Db       1261 TCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTA 1320

Qy       3332 CTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACC 3391

Db       1321 CTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACC 1380

Qy       3392 TCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTG 3451

Db       1381 TCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTG 1440

Qy       3452 TTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATT 3511

Db       1441 TTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATT 1500

Qy       3512 TTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCA 3571

Db       1501 TTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCA 1560

Qy       3572 AAATTCAGGT 3581

Db       1561 AAATTCAGAT 1570
```

RESULT 7
US-09-245-041-10
; Sequence 10, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 1051
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-10

Query Match 12.9%; Score 463; DB 3; Length 1051;
Best Local Similarity 85.2%; Pred. No. 9e-130;
Matches 517; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

```
Qy         15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74

Db        441 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAA 500

Qy         75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134

Db        501 GGAATGTGACCGGCCGTGTGTCAACGGCGGCCGCTGCAACCCTGGCACCGGCCAGTGCGT 560

Qy        135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194

Db        561 CTGCCCCACGGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 620

Qy        195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254

Db        621 TGGCTCTTCTGGATTTGTAACAGATGGACCTGGGAATTATAAATATAAGACGAAGTGCAC 680

Qy        255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314

Db        681 ATGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGCTTCAACCATTTTGCTAC 740

Qy        315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374

Db        741 AGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATCTACGCACCTCTGAT 800

Qy        375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434

Db        801 TGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCTCCTGAGGT 860

Qy        435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494

Db        861 CACTGTCACTTCAGGTTATGCACTGCTGCATTTTTTCAGTGATGCTGCTTATAATCTGAC 920

Qy        495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554

Db        921 TGGATTTAATATCACTTACAATTTTGACATGTGTCCGAATAATTGCTCAGGCCGAGGAGA 980

Qy        555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614

Db        981 GTGTAAGAGCAGTAACAGCAGCAGCGCTGTTGAGTGTGAATGTTCTGAAAACTGGAAAGG 1040

Qy        615 TGAAGCA 621

Db       1041 GGCCGGA 1047
```

RESULT 8
US-09-702-705-966/c
; Sequence 966, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-966

Query Match 6.8%; Score 246; DB 4; Length 246;
Best Local Similarity 100.0%; Pred. No. 2.1e-64;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2293 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 2352

Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy       2353 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 2412
```

```
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy       2413 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 2472
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67

Qy       2473 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 2532
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy       2533 GAAAGG 2538
             ||||||
Db          6 GAAAGG 1


RESULT 9
US-09-736-457-966/c
; Sequence 966, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Wang, Aijun
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.478C15
;  CURRENT APPLICATION NUMBER: US/09/736,457
;  CURRENT FILING DATE:  2000-12-13
;  NUMBER OF SEQ ID NOS:  1864
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 966
;   LENGTH: 246
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-736-457-966

  Query Match             6.8%;  Score 246;  DB 4;  Length 246;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-64;
  Matches  246;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2293 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 2352
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy       2353 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 2412
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy       2413 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 2472
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67

Qy       2473 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 2532
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy       2533 GAAAGG 2538
             ||||||
Db          6 GAAAGG 1


RESULT 10
US-09-614-124B-966/c
; Sequence 966, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
;  APPLICANT: Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.478C9
;  CURRENT APPLICATION NUMBER: US/09/614,124B
;  CURRENT FILING DATE:  2001-07-11
;  NUMBER OF SEQ ID NOS: 1668
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 966
;   LENGTH: 246
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-614-124B-966

  Query Match             6.8%;  Score 246;  DB 4;  Length 246;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-64;
  Matches  246;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2293 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 2352
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy       2353 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 2412
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy       2413 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 2472
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67

Qy       2473 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 2532
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy       2533 GAAAGG 2538
             ||||||
Db          6 GAAAGG 1


RESULT 11
US-09-671-325-966/c
; Sequence 966, Application US/09671325
; Pateht No. 6667154
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Fan, Liqun
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.478C12
;  CURRENT APPLICATION NUMBER: US/09/671,325
;  CURRENT FILING DATE:  2000-09-26
;  NUMBER OF SEQ ID NOS:  1825
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 966
;   LENGTH: 246
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-671-325-966
```

```
  Query Match             6.8%;  Score 246;  DB 4;  Length 246;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-64;
  Matches  246;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2293 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 2352
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy       2353 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 2412
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy       2413 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAAATTTTATCAGAA 2472
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAAATTTTATCAGAA 67

Qy       2473 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 2532
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy       2533 GAAAGG 2538
             ||||||
Db             6 GAAAGG 1


RESULT 12
US-09-833-381-1916
; Sequence 1916, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
;  APPLICANT: Robison, Keith E.
;  TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
;  FILE REFERENCE: 5800-119
;  CURRENT APPLICATION NUMBER: US/09/833,381
;  CURRENT FILING DATE:  2001-04-11
;  PRIOR APPLICATION NUMBER: 09/516,448
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 2050
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1916
;   LENGTH: 415
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(415)
;   OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1916

  Query Match             5.9%;  Score 212.8;  DB 4;  Length 415;
  Best Local Similarity   73.6%;  Pred. No. 3.7e-54;
  Matches  271;  Conservative    0;  Mismatches   97;  Indels    0;  Gaps    0;

Qy        200 CTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGGC 259
              | || || | | | |||||||| ||     || |||||||| ||||| || || || ||||
Db          5 CGTCCGGATATTTAACAGATGGCCCAATTAACTATAAATATAAAACTAAATGTACTTGGC 64

Qy        260 TCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAGT 319
              |||||||||  | ||||||  | |  | ||| |  | ||||||||||||||||||| |
Db         65 TCATTGAAGGCTATCCAAATGCAGTGTTAAGATTAAGATTCAATCATTTTGCTACAGAAT 124

Qy        320 GTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCTG 379
              |||| ||||| ||| | |||||||||||||| || ||||| |||||||  || | ||||
Db        125 GTAGCTGGGATCATATGTATGTTTATGATGGAGATTCAATATATGCACCTTTAATAGCTG 184

Qy        380 CATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTG 439
              | |||||||  | || || ||||| | |   |||||||| ||||| ||||| ||||||
Db        185 TACTTAGTGGTTTGATAGTCCCTGAAATAAGGGGCAATGAAACTGTGCCTGAAGTTGTTA 244

Qy        440 CCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGAT 499
              | ||||| || |||||  || | |||||||||||||||||||| |||||| | ||||| |
Db        245 CTACATCTGGCTATGCACTGTTACATTTTTTTAGTGATGCTGCGTATAATCTAACTGGTT 304

Qy        500 TTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGTA 559
              | || |||   ||      |  ||   ||||| || || ||||| || |  ||  ||||||
Db        305 TCAACATTTTCTATTCAATCAATTCTTGTCCTAACAATTGCTCTGGTCATGGGAAGTGTA 364

Qy        560 AGATCAGT 567
                |  |||
Db        365 CAACTAGT 372


RESULT 13
US-09-245-041-5
; Sequence 5, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT: Nagle, D.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
;  TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
;  FILE REFERENCE: 7853-136
;  CURRENT APPLICATION NUMBER: US/09/245,041
;  CURRENT FILING DATE: 1999-02-05
;  EARLIER APPLICATION NUMBER: 60/093,630
;  EARLIER FILING DATE: 1998-07-21
;  EARLIER APPLICATION NUMBER: 60/104,978
;  EARLIER FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 131
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5
;   LENGTH: 90050
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-245-041-5

  Query Match             5.6%;  Score 202.8;  DB 3;  Length 90050;
  Best Local Similarity   81.8%;  Pred. No. 2.2e-49;
  Matches  234;  Conservative    0;  Mismatches   52;  Indels    0;  Gaps    0;

Qy         947 TTGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAA 1006
               | |||  | |    |  | |    | || |  |||||||||| ||||||||||||||||
Db        6931 TGGTTTTAAAAATCCTGTTTACATCTTTCTTATAGGATAAAATCTACATGTATGGAGGAA 6990

Qy        1007 AAATTGATTCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGT 1066
               |||||||||||| ||||| |||||||||||| ||||||| ||||| |||||||||||  |
Db        6991 AAATTGATTCAACAGGGAACGTGACCAATGAGCTGAGAGTATTTCATATTCATAATGAAT 7050

Qy        1067 CATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCAC 1126
               ||||||| ||||| || || || || |||||  |||||||||||||||||| ||||| ||||
Db        7051 CATGGGTATTGTTAACTCCGAAAGCTAAGGATCAGTATGCAGTGGTTGGACACTCAGCAC 7110

Qy        1127 ACATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTC 1186
               |||||||||||||     |||||| |||||||| |||||||| ||||| ||||| |
Db        7111 ACATTGTTACACTGGCATCTGGCCGTGTGGTCATGTTGGTCATCTTCGGTCATTGCCCAC 7170

Qy        1187 TCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAA 1232
               |||||||||||||||||  |||||||||||||| |||| || | |
Db        7171 TCTATGGATATATAAGCGTTGTGCAGGAATATGACTTGGGTATGTA 7216


RESULT 14
US-09-245-041-3/c
; Sequence 3, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT: Nagle, D.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
;  TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
;  FILE REFERENCE: 7853-136
;  CURRENT APPLICATION NUMBER: US/09/245,041
;  CURRENT FILING DATE: 1999-02-05
;  EARLIER APPLICATION NUMBER: 60/093,630
;  EARLIER FILING DATE: 1998-07-21
;  EARLIER APPLICATION NUMBER: 60/104,978
```

```
;  EARLIER FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 131
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 17056
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-245-041-3

  Query Match             4.7%;  Score 169.2;  DB 3;  Length 17056;
  Best Local Similarity   86.9%;  Pred. No. 9.9e-40;
  Matches  186;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy        510 TTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAA 569
              ||  |||||||| |||||||| ||||| |||||||||||||||||||||||| ||||||
Db       4090 TTGTAGTTTTGACATGTGTCCGAATAATTGCTCAGGCCGAGGAGAGTGTAAGAGCAGTAA 4031

Qy        570 TAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACAT 629
              ||||||    ||||||| |||||||||||||||||||||||| ||  | ||||||||
Db       4030 CAGCAGCAGCGCTGTTGAGTGTGAATGTTCTGAAAACTGGAAAGGGGAGTCGTGTGACAT 3971

Qy        630 TCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGT 689
              |||||||||||||||||||||||| |||||||| |||||||||| ||| |||| |||
Db       3970 TCCTCACTGTACAGACAACTGTGGCTTTCCTCACCGAGGCATCTGTAATGCAAGCGATAC 3911

Qy        690 CAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGT 723
              |||||| ||||||||||   |  ||||||||||
Db       3910 CAGAGGGTGCTCCTGCTTTCCTCACTGGCAGGGT 3877


RESULT 15
US-09-245-041-6/c
; Sequence 6, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT: Nagle, D.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
;  TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
;  FILE REFERENCE: 7853-136
;  CURRENT APPLICATION NUMBER: US/09/245,041
;  CURRENT FILING DATE: 1999-02-05
;  EARLIER APPLICATION NUMBER: 60/093,630
;  EARLIER FILING DATE: 1998-07-21
;  EARLIER APPLICATION NUMBER: 60/104,978
;  EARLIER FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 131
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 6
;   LENGTH: 226
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-245-041-6

  Query Match             3.7%;  Score 132.2;  DB 3;  Length 226;
  Best Local Similarity   78.6%;  Pred. No. 6.7e-30;
  Matches  158;  Conservative    0;  Mismatches   43;  Indels    0;  Gaps    0;

Qy       2389 GATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCT 2448
              |||||||| ||||| |||||||||| || |||||||||| ||||||||||||||||||
Db        225 GATATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGTGATGCT 166

Qy       2449 GGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATC 2508
              || |||||||| || || ||||| || |||||||||||| | ||||||||||||||||||
Db        165 GGCTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGATTAAAGGCTGCAACCTGCATC 106

Qy       2509 AACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGG 2568
              ||||| |||||||| || ||||||||||||||  ||  |||  | |    |  |
Db        105 AACCCTCTCAATGGCAGCGTCTGTGAAAGGCCTGGTAAGGACATGGGTGCATATAGTGCT 46

Qy       2569 ACACCATGTGCCTTGAGGACA 2589
              ||  | | |||  ||   ||
Db         45 CCAGGAGGAGCCAAGACAGCA 25
```

```
Search completed: March  5, 2004, 12:20:03
Job time : 175.851 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2004, 16:07:54 ; Search time 10236 Seconds
(without alignments)
17225.392 Million cell updates/sec

Title: US-09-787-097-11
Perfect score: 4068
Sequence: 1 atggtggccgcagcggcggc.....agcctgggacctgcatctga 4068

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3988.4	98.0	8589	6	AR164816	AR164816 Sequence
2	3566	87.7	4493	4	AF531101	AF531101 Bos tauru
3	3556.4	87.4	3597	9	AF034957	AF034957 Homo sapi
4	3499.8	86.0	4072	6	AR164817	AR164817 Sequence
5	3492.4	85.9	3819	9	AF106861	AF106861 Homo sapi
6	3395.4	83.5	4284	10	AB062913	AB062913 Mesocrice
7	3377.4	83.0	6083	10	AF116897	AF116897 Mus muscu
8	3354.8	82.5	4313	10	AF119821	AF119821 Mus muscu
9	3353.2	82.4	8739	10	AB038387	AB038387 Rattus no
10	3326.4	81.8	8827	6	AR164807	AR164807 Sequence
11	2928.8	72.0	4500	10	AB038388	AB038388 Rattus no
12	2694.6	66.2	5683	10	AB093245	AB093245 Mus muscu
13	2051.8	50.4	2625	6	AR164818	AR164818 Sequence
14	1853	45.6	6370	6	AR164815	AR164815 Sequence
15	1697	41.7	2578	9	AK000356	AK000356 Homo sapi
16	1466.6	36.1	4463	9	AK127277	AK127277 Homo sapi
17	1465	36.0	4140	9	AY442317	AY442317 Homo sapi
18	1451.8	35.7	2419	6	AR164813	AR164813 Sequence
19	1361	33.5	5632	9	AB011120	AB011120 Homo sapi
20	1141.6	28.1	7302	9	AB011106	AB011106 Homo sapi
21	1131	27.8	3609	6	AX503776	AX503776 Sequence
22	1098	27.0	5002	10	BC027764	BC027764 Mus muscu
23	817.8	20.1	1012	6	E26747	E26747 Structure a
24	704.2	17.3	4997	10	BC050020	BC050020 Mus muscu
25	585.8	14.4	3297	10	BC030872	BC030872 Mus muscu
26	543	13.3	543	6	AX778545	AX778545 Sequence
27	463	11.4	1051	6	AR164814	AR164814 Sequence
28	418.4	10.3	1818	9	BC047716	BC047716 Homo sapi
c 29	332.2	8.2	110000	2	BX004762_0	BX004762 Mus muscu
c 30	332.2	8.2	110000	2	BX004762_1	Continuation (2 of
31	332.2	8.2	110261	10	AL807401	AL807401 Mouse DNA
32	332.2	8.2	335799	2	BX005139	BX005139 Mus muscu
33	297.4	7.3	315	4	AF301008	AF301008 Bos tauru
34	289.4	7.1	110000	2	AL954766_2	Continuation (3 of
35	289.4	7.1	110000	2	BX004762_3	Continuation (4 of
36	253.8	6.2	982	6	AX364943	AX364943 Sequence
37	250.8	6.2	2018	9	H009384S08	AF218896 Homo sapi
38	250.8	6.2	104907	9	HSDJ741H3	AL132773 Human DNA
c 39	250.8	6.2	155869	2	AC023420	AC023420 Homo sapi
c 40	250.8	6.2	167357	2	AC055771	AC055771 Homo sapi
41	250.8	6.2	206647	9	AP002898	AP002898 Homo sapi
c 42	246	6.0	246	6	AR273223	AR273223 Sequence
c 43	246	6.0	246	6	AR276804	AR276804 Sequence
c 44	246	6.0	246	6	AR407079	AR407079 Sequence
c 45	246	6.0	246	6	AX368256	AX368256 Sequence

ALIGNMENTS

RESULT 1
AR164816
LOCUS AR164816 8589 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 14 from patent US 6274339.
ACCESSION AR164816
VERSION AR164816.1 GI:16238025
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 8589)
AUTHORS Moore,K. and Nagle,D.Lynn.
TITLE Methods and compositions for the diagnosis and treatment of body weight disorders, including obesity
JOURNAL Patent: US 6274339-A 14 14-AUG-2001;

```
FEATURES             Location/Qualifiers
     source          1. .8589
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             98.0%;  Score 3988.4;  DB 6;  Length 8589;
  Best Local Similarity   99.0%;  Pred. No. 0;
  Matches 4013;  Conservative    0;  Mismatches   41;  Indels    0;  Gaps    0;

Qy         15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74
              || |||    |  ||   | ||  |  |  |  | | |   |||| ||   || || ||
Db        234 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAA 293

Qy         75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134
              |   | |  | |||  ||||||||||||||||||||||||||||||||||||||||||
Db        294 GGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 353

Qy        135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        354 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 413

Qy        195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 473

Qy        255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 533

Qy        315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 593

Qy        375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 653

Qy        435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        654 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 713

Qy        495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        714 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 773

Qy        555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        774 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 833

Qy        615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        834 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 893

Qy        675 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        894 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 953

Qy        735 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        954 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 1013

Qy        795 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1014 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 1073

Qy        855 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1074 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 1133

Qy        915 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1134 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 1193

Qy        975 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1194 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1253

Qy       1035 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1254 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1313

Qy       1095 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1314 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1373

Qy       1155 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1374 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1433

Qy       1215 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1434 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1493

Qy       1275 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1494 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1553

Qy       1335 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1554 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1613

Qy       1395 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1454
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1614 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1673

Qy       1455 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1674 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1733

Qy       1515 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1734 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1793

Qy       1575 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGC 1634
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1794 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGC 1853

Qy       1635 AGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGA 1694
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1854 AGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGA 1913

Qy       1695 CATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGC 1754
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1914 CATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGC 1973

Qy       1755 AGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGC 1814
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1974 AGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGC 2033

Qy       1815 GCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCT 1874
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2034 GCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCT 2093

Qy       1875 TGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGA 1934
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2094 TGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGA 2153

Qy       1935 CTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGAT 1994
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2154 CTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGAT 2213

Qy       1995 CTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAA 2054
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2214 CTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAA 2273
```

```
Qy       2055 GACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGA 2114
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2274 GACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGA 2333

Qy       2115 GTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATG 2174
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2334 GTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATG 2393

Qy       2175 TTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCA 2234
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2394 TTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCA 2453

Qy       2235 CAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCT 2294
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2454 CAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCT 2513

Qy       2295 GCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCT 2354
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2514 GCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCT 2573

Qy       2355 TCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTT 2414
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2574 TCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTT 2633

Qy       2415 ACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACC 2474
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2634 ACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACC 2693

Qy       2475 CAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGA 2534
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2694 CAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGA 2753

Qy       2535 AAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATG 2594
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2754 AAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATG 2813

Qy       2595 TGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGT 2654
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2814 TGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGT 2873

Qy       2655 GGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGAT 2714
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2874 GGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGAT 2933

Qy       2715 GAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCA 2774
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2934 GAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCA 2993

Qy       2775 ACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGG 2834
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2994 ACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGG 3053

Qy       2835 TTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACA 2894
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3054 TTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACA 3113

Qy       2895 GCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCA 2954
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3114 GCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCA 3173

Qy       2955 CTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAA 3014
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3174 CTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAA 3233

Qy       3015 GTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGA 3074
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3234 GTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGA 3293

Qy       3075 TCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAA 3134
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3294 TCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAA 3353

Qy       3135 CACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCT 3194
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3354 CACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCT 3413

Qy       3195 ATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCT 3254
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3414 ATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCT 3473

Qy       3255 TCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGC 3314
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3474 TCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGC 3533

Qy       3315 TATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGC 3374
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3534 TATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGC 3593

Qy       3375 CTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGC 3434
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3594 CTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGC 3653

Qy       3435 TGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTC 3494
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3654 TGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTC 3713

Qy       3495 TAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAA 3554
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3714 TAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAA 3773

Qy       3555 TTTCACCTGGCCCATCAAAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCT 3614
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3774 TTTCACCTGGCCCATCAAAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCT 3833

Qy       3615 GGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGT 3674
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3834 GGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGT 3893

Qy       3675 GGTTTGGAAGATCAAACAAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGA 3734
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3894 GGTTTGGAAGATCAAACAAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGA 3953

Qy       3735 GATGCAACAGATGGCCAGCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGA 3794
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3954 GATGCAACAGATGGCCAGCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGA 4013

Qy       3795 GGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGA 3854
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4014 GGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGA 4073

Qy       3855 GCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCT 3914
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4074 GCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCT 4133

Qy       3915 GGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACAT 3974
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4134 GGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACAT 4193

Qy       3975 TTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCA 4034
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4194 TTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCA 4253

Qy       4035 GCAGCCCCCTGCACAGCCTGGGACCTGCATCTGA 4068
             ||||||||||||||||||||||||||||||||||
Db       4254 GCAGCCCCCTGCACAGCCTGGGACCTGCATCTGA 4287
```

```
RESULT 2
AF531101
LOCUS       AF531101                4493 bp    mRNA    linear   MAM 19-AUG-2002
DEFINITION  Bos taurus attractin mRNA, complete cds.
ACCESSION   AF531101
VERSION     AF531101.1  GI:22297307
KEYWORDS    .
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 4493)
  AUTHORS   Graphodatskaya,D., Joerg,H. and Stranzinger,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2002) Animal Sciences, ETH, Tannenstrasse 1,
            Zurich, CH 8092, Switzerland
FEATURES             Location/Qualifiers
     source          1. .4493
                     /organism="Bos taurus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9913"
                     /chromosome="13"
     CDS             52. .4299
                     /note="transmembrane form; mahogany protein"
                     /codon_start=1
                     /product="attractin"
                     /protein_id="AAM95445.1"
                     /db_xref="GI:22297308"
                     /translation="MVAAATEARLRRTVVTPAPTSRNGRRCRDWDATRAGRPGLRTGL
                     RLPRRWRRLLLLLLPPLLLLPWAAEAAAAAASVSGSAAAEAKECDRPCVNGGRCNPGT
                     GQCVCPAGWVGEQCQHCGGRFRLTGSSGFITDGPGNYKYKTKCTWLIEGQPNRIMRLR
                     FNHFATECSWDHLYVYDGDSIYAPLVAAFSGLIVPERDGNETVPEVVATSGYALLHFF
                     SDAAYNLTGFNITYNFDMCPNNCSGRGECKISNSSNTVQCECSENWKGEACDIPHCVN
                     NCGFPHRGICNSSDVRGCSCFSEWQGPGCSVPVPANQSFWTREEYSDLKLPRASHKAV
                     VNGNIMWVVGGYMFNHSDYNMVLAYDLASREWLALNRSVNSVVVRYGHSLALYQDKIY
                     MYGGKIDSTGNVTNELRVFHIHNESWVLLSPKAKEQYAVVGHSAHIVTLKNGRVVMLV
                     IFGHCPLYGYISSVQEYDLDKNTWSILHTQGALVQGGYGHSSVYDDRTKALYIHGGYK
                     AFSANKYRLADDLYRYDVDTQMWTILKDSRFFRYLHTAVIVSGTMLVFGGNTHNDTSM
                     SHGAKCFSSDFMAYDIDCDRWSVLPRPDLHHDVNRFGHSAVLYNRTMYVFGGFNSLLL
                     SDILVFTSEQCEAHQSEAACLAAGPGVRCMWDTGSSQCVSWELAPEAQEKIKSECFSK
                     RIFDHDRCDQLTDCYSCTANTNGCQWCNDHCVPTNHSCTEGQISIFKYDHCPKDNPMY
                     YCNKKTSCRSCALDQNCQWEPRNQECIALPENICGIGWHLVGNSCLKITTARETYDNA
                     KLSCRNHNAFLASLTTQKKVEFVLKQLRIMQSSQSMSKFTLTPWVGLRKINVSYWCWE
                     DMSPFTNSSLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLNGSVCERPANHSAKQ
                     CRTPCALRTACGECTSGSSECMWCSNMKQCVDSNAYVASFPFGQCMEWYTMSSCPPEN
                     CSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKMPSQGPTGNSYPQPLLNS
                     SMCLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCENLTTGKHCETCISGFYGDPTN
                     GGKCQPCRCNGHASLCNTNTGKCFCTTKGVKGDECQLCEVENRYQGNPLKGTCYYTLL
                     IDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASKNFNLNITWAASFSAGTQ
                     AGEEMPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTWPIKIQIAFSQHSNF
                     MDLVQFFVTFFSCFLSLLLVAAVVWKIKQSCWASRRREQLLREMQQMASRPFASVNVA
                     LETDEEPPDLIGGSIKTVPKPIALEPCFGNKAAVLSVFVRLPRGLGGIPPPGQSGLAV
                     ASALVDISQQMPVVYKEKSGAVRNRKQQPPAQPGTCI"
ORIGIN

  Query Match           87.7%;  Score 3566;  DB 4;  Length 4493;
  Best Local Similarity  92.6%;  Pred. No. 0;
  Matches 3756;  Conservative    0;  Mismatches  295;  Indels    3;  Gaps    1;

Qy         15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74
Db        249 GGCGGCTGAGGCCGCGGCGGCAGCGGCGTCGGTGTCGGGCTCAGCCGCAGCCGAGGCCAA 308

Qy         75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134
Db        309 GGAATGTGACCGGCCCTGTGTCAACGGCGGCCGCTGCAACCCTGGCACCGGCCAGTGTGT 368

Qy        135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194
Db        369 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 428

Qy        195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254
Db        429 TGGATCTTCTGGATTTATAACAGATGGACCTGGAAATTATAAATATAAAACAAAATGCAC 488

Qy        255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314
Db        489 ATGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 548

Qy        315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374
Db        549 AGAGTGTAGTTGGGACCATTTATATGTTTATGACGGGGACTCAATTTATGCACCCCTAGT 608

Qy        375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434
Db        609 TGCTGCCTTTAGCGGCCTCATTGTTCCTGAGAGAGACGGCAATGAGACTGTCCCAGAGGT 668

Qy        435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494
Db        669 TGTCGCCACATCAGGTTATGCCCTGCTGCATTTTTTTAGTGATGCTGCTTATAACTTGAC 728

Qy        495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554
Db        729 TGGATTTAATATCACTTACAATTTTGACATGTGTCCAAATAATTGCTCAGGCCGAGGAGA 788

Qy        555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614
Db        789 ATGTAAGATCAGTAATAGCAGCAACACTGTTCAATGTGAATGTTCTGAAAACTGGAAAGG 848

Qy        615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 674
Db        849 TGAAGCATGTGACATTCCTCACTGTGTAAATAACTGTGGTTTTCCTCATCGAGGCATCTG 908

Qy        675 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 734
Db        909 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGAGTGGCAAGGTCCTGGATGTTC 968

Qy        735 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 794
Db        969 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACACGAGAGGAATATTCTGACCTAAAGCT 1028

Qy        795 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 854
Db       1029 CCCCAGAGCCTCTCACAAAGCTGTGGTCAATGGAAATATAATGTGGGTTGTTGGCGGATA 1088

Qy        855 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 914
Db       1089 TATGTTCAACCACTCAGATTACAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 1148

Qy        915 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 974
Db       1149 GCTTGCACTGAACCGTTCTGTGAACAGTGTGGTTGTTAGATATGGCCATTCTTTAGCATT 1208

Qy        975 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1034
Db       1209 ATACCAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACAGGGAATGTGACCAA 1268

Qy       1035 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1094
Db       1269 TGAGTTGAGAGTGTTTCATATTCATAATGAATCATGGGTGTTGTTGTCCCCCAAAGCAAA 1328

Qy       1095 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1154
Db       1329 GGAGCAGTATGCAGTGGTTGGGCACTCGGCGCACATTGTCACGTTGAAAAATGGCCGAGT 1388

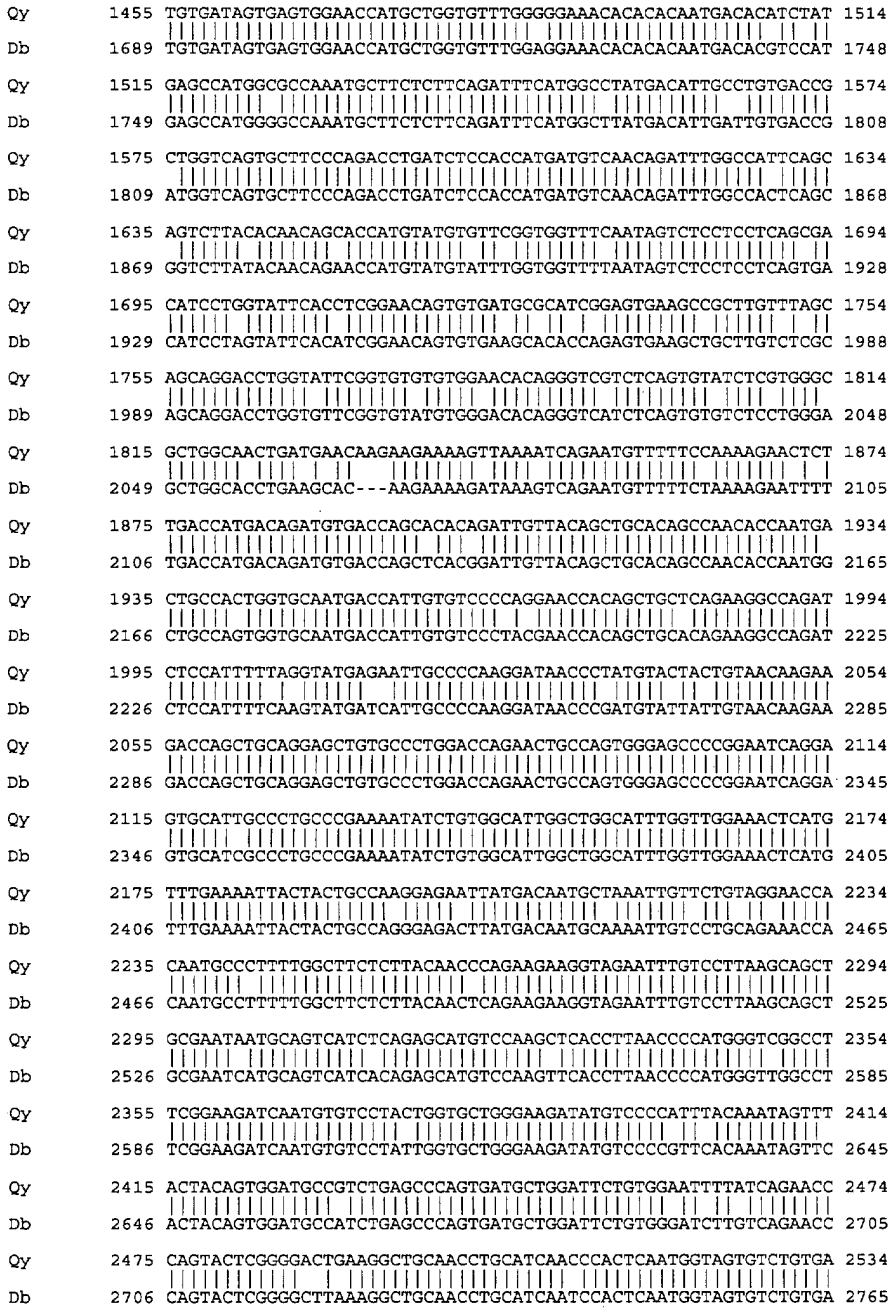
Qy       1155 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1214
Db       1389 GGTCATGCTGGTTATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAGTGTGCAGGA 1448

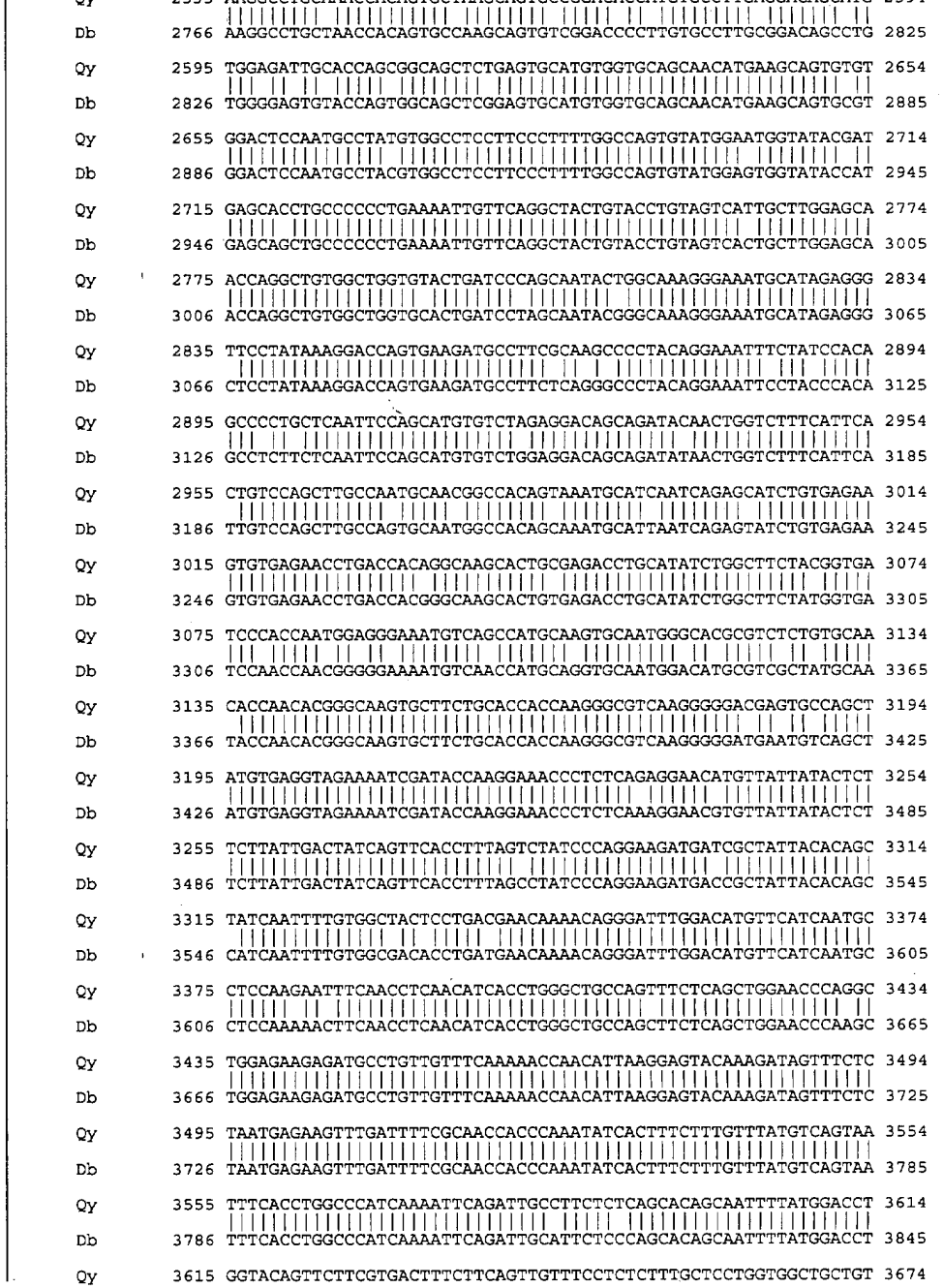
Qy       1215 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1274
Db       1449 ATATGACTTGGATAAGAACACATGGAGTATATTACACACTCAGGGTGCCCTTGTCCAAGG 1508

Qy       1275 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1334
Db       1509 CGGCTACGGCCATAGCAGTGTCTATGATGATAGGACCAAGGCCCTGTATATTCACGGAGG 1568

Qy       1335 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1394
Db       1569 CTACAAGGCTTTCAGTGCCAATAAATACCGGCTTGCAGATGATCTCTACAGATATGATGT 1628

Qy       1395 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1454
Db       1629 GGACACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1688
```


Qy 1455 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1514
Db 1689 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGAGGAAACACACACAATGACACGTCCAT 1748
Qy 1515 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1574
Db 1749 GAGCCATGGGGCCAAATGCTTCTCTTCAGATTTCATGGCTTATGACATTGATTGTGACCG 1808
Qy 1575 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGC 1634
Db 1809 ATGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCACTCAGC 1868
Qy 1635 AGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGA 1694
Db 1869 GGTCTTATACAACAGAACCATGTATGTATTTGGTGGTTTTAATAGTCTCCTCCTCAGTGA 1928
Qy 1695 CATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGC 1754
Db 1929 CATCCTAGTATTCACATCGGAACAGTGTGAAGCACACCAGAGTGAAGCTGCTTGTCTCGC 1988
Qy 1755 AGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGC 1814
Db 1989 AGCAGGACCTGGTGTTCGGTGTATGTGGGACACAGGGTCATCTCAGTGTGTCTCCTGGGA 2048
Qy 1815 GCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCT 1874
Db 2049 GCTGGCACCTGAAGCAC---AAGAAAAGATAAAGTCAGAATGTTTTTCTAAAAGAATTTT 2105
Qy 1875 TGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGA 1934
Db 2106 TGACCATGACAGATGTGACCAGCTCACGGATTGTTACAGCTGCACAGCCAACACCAATGG 2165
Qy 1935 CTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGAT 1994
Db 2166 CTGCCAGTGGTGCAATGACCATTGTGTCCCTACGAACCACAGCTGCACAGAAGGCCAGAT 2225
Qy 1995 CTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAA 2054
Db 2226 CTCCATTTTCAAGTATGATCATTGCCCCAAGGATAACCCGATGTATTATTGTAACAAGAA 2285
Qy 2055 GACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGA 2114
Db 2286 GACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGA 2345
Qy 2115 GTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATG 2174
Db 2346 GTGCATCGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATG 2405
Qy 2175 TTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCA 2234
Db 2406 TTTGAAAATTACTACTGCCAGGGAGACTTATGACAATGCAAAATTGTCCTGCAGAAACCA 2465
Qy 2235 CAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCT 2294
Db 2466 CAATGCCTTTTTGGCTTCTCTTACAACTCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCT 2525
Qy 2295 GCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCT 2354
Db 2526 GCGAATCATGCAGTCATCACAGAGCATGTCCAAGTTCACCTTAACCCCATGGGTTGGCCT 2585
Qy 2355 TCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTT 2414
Db 2586 TCGGAAGATCAATGTGTCCTATTGGTGCTGGGAAGATATGTCCCCGTTCACAAATAGTTC 2645
Qy 2415 ACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACC 2474
Db 2646 ACTACAGTGGATGCCATCTGAGCCCAGTGATGCTGGATTCTGTGGGATCTTGTCAGAACC 2705
Qy 2475 CAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGA 2534
Db 2706 CAGTACTCGGGGCTTAAAGGCTGCAACCTGCATCAATCCACTCAATGGTAGTGTCTGTGA 2765


Qy 2535 AAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATG 2594
Db 2766 AAGGCCTGCTAACCACAGTGCCAAGCAGTGTCGGACCCCTTGTGCCTTGCGGACAGCCTG 2825
Qy 2595 TGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGT 2654
Db 2826 TGGGGAGTGTACCAGTGGCAGCTCGGAGTGCATGTGGTGCAGCAACATGAAGCAGTGCGT 2885
Qy 2655 GGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGAT 2714
Db 2886 GGACTCCAATGCCTACGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAGTGGTATACCAT 2945
Qy 2715 GAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCA 2774
Db 2946 GAGCAGCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCACTGCTTGGAGCA 3005
Qy 2775 ACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGG 2834
Db 3006 ACCAGGCTGTGGCTGGTGCACTGATCCTAGCAATACGGGCAAAGGGAAATGCATAGAGGG 3065
Qy 2835 TTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACA 2894
Db 3066 CTCCTATAAAGGACCAGTGAAGATGCCTTCTCAGGGCCCTACAGGAAATTCCTACCCACA 3125
Qy 2895 GCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCA 2954
Db 3126 GCCTCTTCTCAATTCCAGCATGTGTCTGGAGGACAGCAGATATAACTGGTCTTTCATTCA 3185
Qy 2955 CTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAA 3014
Db 3186 TTGTCCAGCTTGCCAGTGCAATGGCCACAGCAAATGCATTAATCAGAGTATCTGTGAGAA 3245
Qy 3015 GTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGA 3074
Db 3246 GTGTGAGAACCTGACCACGGGCAAGCACTGTGAGACCTGCATATCTGGCTTCTATGGTGA 3305
Qy 3075 TCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAA 3134
Db 3306 TCCAACCAACGGGGGAAAATGTCAACCATGCAGGTGCAATGGACATGCGTCGCTATGCAA 3365
Qy 3135 CACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCT 3194
Db 3366 TACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGATGAATGTCAGCT 3425
Qy 3195 ATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCT 3254
Db 3426 ATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAAAGGAACGTGTTATTATACTCT 3485
Qy 3255 TCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGC 3314
Db 3486 TCTTATTGACTATCAGTTCACCTTTAGCCTATCCCAGGAAGATGACCGCTATTACACAGC 3545
Qy 3315 TATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGC 3374
Db 3546 CATCAATTTTGTGGCGACACCTGATGAACAAAACAGGGATTTGGACATGTTCATCAATGC 3605
Qy 3375 CTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGC 3434
Db 3606 CTCCAAAAACTTCAACCTCAACATCACCTGGGCTGCCAGCTTCTCAGCTGGAACCCAAGC 3665
Qy 3435 TGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTC 3494
Db 3666 TGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTC 3725
Qy 3495 TAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAA 3554
Db 3726 TAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAA 3785
Qy 3555 TTTCACCTGGCCCATCAAAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCT 3614
Db 3786 TTTCACCTGGCCCATCAAAATTCAGATTGCATTCTCCCAGCACAGCAATTTTATGGACCT 3845
Qy 3615 GGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGT 3674

```
Db       3846 AGTACAGTTCTTCGTGACCTTCTTCAGTTGTTTCCTCTCTCTGCTCCTGGTGGCTGCTGT 3905

Qy       3675 GGTTTGGAAGATCAAACAAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGA 3734
             |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db       3906 GGTTTGGAAGATCAAACAAAGTTGTTGGGCTTCCAGACGTAGAGAGCAACTTCTTCGAGA 3965

Qy       3735 GATGCAACAGATGGCCAGCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGA 3794
             |||||| |||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db       3966 GATGCAGCAGATGGCCAGTCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGA 4025

Qy       3795 GGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGA 3854
             |||||||||||||||||||||||||||||||||||||||||||||| |||||||| |||||
Db       4026 AGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAGCCCATTGCCCTGGA 4085

Qy       3855 GCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCT 3914
             |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db       4086 GCCGTGTTTTGGCAACAAAGCTGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCT 4145

Qy       3915 GGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACAT 3974
             ||||| |||||||||||||||||| || || || |||||||| ||||||||||||||||
Db       4146 CGGTGGAATCCCTCCTCCTGGGCAGTCCGGGCTCGCAGTGGCCAGTGCCCTGGTGGACAT 4205

Qy       3975 TTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCA 4034
             |||||||||||||| | || ||||| |||||||||||||| |||||||||| || ||
Db       4206 TTCTCAGCAGATGCCAGTCGTCTACAAAGAGAAGTCAGGAGCAGTGAGAAACCGAAAACA 4265

Qy       4035 GCAGCCCCCTGCACAGCCTGGGACCTGCATCTGA 4068
             |||||||||||| |||||||||||||||| |||
Db       4266 GCAGCCCCCTGCGCAGCCTGGGACCTGCATTTGA 4299
```

RESULT 3
AF034957
LOCUS AF034957 3597 bp mRNA linear PRI 30-SEP-1998
DEFINITION Homo sapiens secreted T cell activation protein Attractin (attractin) mRNA, complete cds.
ACCESSION AF034957
VERSION AF034957.1 GI:3676346
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3597)
AUTHORS Duke-Cohan,J.S., Gu,J., McLaughlin,D.F., Xu,Y., Freeman,G.J. and Schlossman,S.F.
TITLE Attractin (DPPT-L), a member of the CUB family of cell adhesion and guidance proteins, is secreted by activated human T lymphocytes and modulates immune cell interactions
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (19), 11336-11341 (1998)
MEDLINE 98409658
PUBMED 9736737
REFERENCE 2 (bases 1 to 3597)
AUTHORS Duke-Cohan,J.S., Gu,J., Ao,Z., McLaughlin,D.F., Freeman,G.J. and Schlossman,S.F.
TITLE Direct Submission
JOURNAL Submitted (19-NOV-1997) Division of Cancer, Immunology and AIDS, Dana Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, USA
FEATURES Location/Qualifiers
source 1. .3597
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
gene 1. .3597
/gene="attractin"
CDS 1. .3597
/gene="attractin"
/note="secreted T cell activation protein; Method: conceptual translation with partial peptide sequencing; DPPT-L"
/codon_start=1
/product="Attractin"
/protein_id="AAC61902.1"
/db_xref="GI:3676347"
/translation="MVAAAAATEARLRRRTAATAALAGRSGGPHCVNGGRCNPGTGQC
VCPAGWVGEQCQHCGGRFRLTGSSGFVTDGPGNYKYKTKCTWLIEGQPNRIMRLRFNH
FATECSWDHLYVYDGDSIYAPLVAAFSGLIVPERDGNETVPEVVATSGYALLHFFSDA
AYNLTGFNITYSFDMCPNNCSGRGECKISNSSDTVECECSENWKGEACDIPHCTDNCG
FPHRGICNSSDVRGCSCFSDWQGPGCSVPVPANQSFWTREEYSNLKLPRASHKAVVNG
NIMWVVGGYMFNHSDYNMVLAYDLASREWLPLNRSVNNVVVRYGHSLALYKDKIYMYG
GKIDSTGNVTNELRVFHIHNESWVLLTPKAKEQYAVVGHSAHIVTLKNGRVVMLVIFG
HCPLYGYISNVQEYDLDKNTWSILHTQGALVQGGYGHSSVYDHRTRALYVHGGYKAFS
ANKYRLADDLYRYDVDTQMWTILKDSRFFRYLHTAVIVSGTMLVFGGNTHNDTSMSHG
AKCFSSDFMAYDIACDRWSVLPRPDSTMMSTDLAIPAVLHNSTMYVFGGFNSLLLSDI
LVFTSEQCDAHRSEAACLAAGPGIRCVWNTGSSQCISWALATDEQEEKLKSECFSKRT
LDHDRCDQHTDCYSCTANTNDCHWCNDHCVPRNHSCSEGQISIFRYENCPKDNPMYYC
NKKTSCRSCALDQNCQWEPRNQECIALPENICGIGWHLVGNSCLKITTAKENYDNAKL
FCRNHNALLASLTTQKKVEFVLKQLRIMQSSQSMSKLTLTPWVGLRKINVSYWCWEDM
SPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLNGSVCERPANHSAKQCR
TPCALRTACGDCTSGSSECMWCSNMKQCVDSNAYVASFPFGQCMEWYTMSTCPPENCS
GYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKMPSQAPTGNFYPQPLLNSSM
CLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCENLTTGKHCETCISGFYGDPTNGG
KCQPCKCNGHASLCNTNTGKCFCTTKGVKGDECQLCEVENRYQGNPLRGTCYYTLLID
YQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASKNFNLNITWAASFSAGTQAG
EEMPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTWPIKIQVQTEQ"

ORIGIN

Query Match 87.4%; Score 3556.4; DB 9; Length 3597;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3579; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

```
Qy          1 ATGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 60

Qy         61 GCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGC 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGC 120

Qy        121 ACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGC 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGC 180

Qy        181 CGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATAC 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATAC 240

Qy        241 AAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTC 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTC 300

Qy        301 AATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATT 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATT 360

Qy        361 TATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAG 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAG 420

Qy        421 ACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCT 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCT 480

Qy        481 GCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGC 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGC 540

Qy        541 TCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCT 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCT 600

Qy        601 GAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCT 660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCT 660
```

```
Qy       661 CATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAG 720
Db       661 CATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAG 720

Qy       721 GGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATAT 780
Db       721 GGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATAT 780

Qy       781 TCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGG 840
Db       781 TCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGG 840

Qy       841 GTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTT 900
Db       841 GTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTT 900

Qy       901 GCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGT 960
Db       901 GCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGT 960

Qy       961 CATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACT 1020
Db       961 CATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACT 1020

Qy      1021 GGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTG 1080
Db      1021 GGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTG 1080

Qy      1081 ACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTG 1140
Db      1081 ACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTG 1140

Qy      1141 AAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATA 1200
Db      1141 AAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATA 1200

Qy      1201 AGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGT 1260
Db      1201 AGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGT 1260

Qy      1261 GCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTA 1320
Db      1261 GCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTA 1320

Qy      1321 TACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTC 1380
Db      1321 TACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTC 1380

Qy      1381 TACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGT 1440
Db      1381 TACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGT 1440

Qy      1441 TACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACAC 1500
Db      1441 TACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACAC 1500

Qy      1501 AATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGAC 1560
Db      1501 AATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGAC 1560

Qy      1561 ATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGA 1620
Db      1561 ATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGAT-TCCACCATGATGTCAACAGA 1619

Qy      1621 TTTGGCCATT-CAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAG 1679
Db      1620 TTTGGCCATTCCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAG 1679

Qy      1680 TCTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGA 1739
Db      1680 TCTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGA 1739
```

```
Qy      1740 AGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCA 1799
Db      1740 AGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCA 1799

Qy      1800 GTGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTT 1859
Db      1800 GTGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTT 1859

Qy      1860 TTCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCAC 1919
Db      1860 TTCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCAC 1919

Qy      1920 AGCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTG 1979
Db      1920 AGCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTG 1979

Qy      1980 CTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTA 2039
Db      1980 CTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTA 2039

Qy      2040 CTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGA 2099
Db      2040 CTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGA 2099

Qy      2100 GCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTT 2159
Db      2100 GCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTT 2159

Qy      2160 GGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATT 2219
Db      2160 GGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATT 2219

Qy      2220 GTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATT 2279
Db      2220 GTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATT 2279

Qy      2280 TGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAAC 2339
Db      2280 TGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAAC 2339

Qy      2340 CCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCC 2399
Db      2340 CCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCC 2399

Qy      2400 ATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGG 2459
Db      2400 ATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGG 2459

Qy      2460 AATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAA 2519
Db      2460 AATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAA 2519

Qy      2520 TGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGC 2579
Db      2520 TGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGC 2579

Qy      2580 CTTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAA 2639
Db      2580 CTTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAA 2639

Qy      2640 CATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTAT 2699
Db      2640 CATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTAT 2699

Qy      2700 GGAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAG 2759
Db      2700 GGAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAG 2759

Qy      2760 TCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGG 2819
Db      2760 TCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGG 2819

Qy      2820 GAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGG 2879
```

```
Db       2820 GAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGG 2879

Qy       2880 AAATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAA 2939
Db       2880 AAATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAA 2939

Qy       2940 CTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCA 2999
Db       2940 CTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCA 2999

Qy       3000 GAGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATC 3059
Db       3000 GAGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATC 3059

Qy       3060 TGGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCA 3119
Db       3060 TGGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCA 3119

Qy       3120 CGCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGG 3179
Db       3120 CGCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGG 3179

Qy       3180 GGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAAC 3239
Db       3180 GGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAAC 3239

Qy       3240 ATGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGA 3299
Db       3240 ATGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGA 3299

Qy       3300 TCGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGA 3359
Db       3300 TCGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGA 3359

Qy       3360 CATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTC 3419
Db       3360 CATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTC 3419

Qy       3420 AGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTA 3479
Db       3420 AGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTA 3479

Qy       3480 CAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTT 3539
Db       3480 CAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTT 3539

Qy       3540 TGTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGAT 3581
Db       3540 TGTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGGT 3581
```

```
RESULT 4
AR164817
LOCUS       AR164817                4072 bp    DNA     linear   PAT 17-OCT-2001
DEFINITION  Sequence 16 from patent US 6274339.
ACCESSION   AR164817
VERSION     AR164817.1  GI:16238027
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 4072)
  AUTHORS   Moore,K. and Nagle,D.Lynn.
  TITLE     Methods and compositions for the diagnosis and treatment of body
            weight disorders, including obesity
  JOURNAL   Patent: US 6274339-A 16 14-AUG-2001;
FEATURES             Location/Qualifiers
     source          1. .4072
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN
```

```
Query Match             86.0%;  Score 3499.8;  DB 6;  Length 4072;
Best Local Similarity   98.8%;  Pred. No. 0;
Matches 3525;  Conservative    0;  Mismatches   42;  Indels    0;  Gaps    0;

Qy        15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGGCTCGCGGGCAG 74
Db       234 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAA 293

Qy        75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134
Db       294 GGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 353

Qy       135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194
Db       354 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 413

Qy       195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254
Db       414 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 473

Qy       255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314
Db       474 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 533

Qy       315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374
Db       534 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 593

Qy       375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434
Db       594 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 653

Qy       435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494
Db       654 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 713

Qy       495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554
Db       714 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 773

Qy       555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614
Db       774 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 833

Qy       615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 674
Db       834 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 893

Qy       675 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 734
Db       894 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 953

Qy       735 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 794
Db       954 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 1013

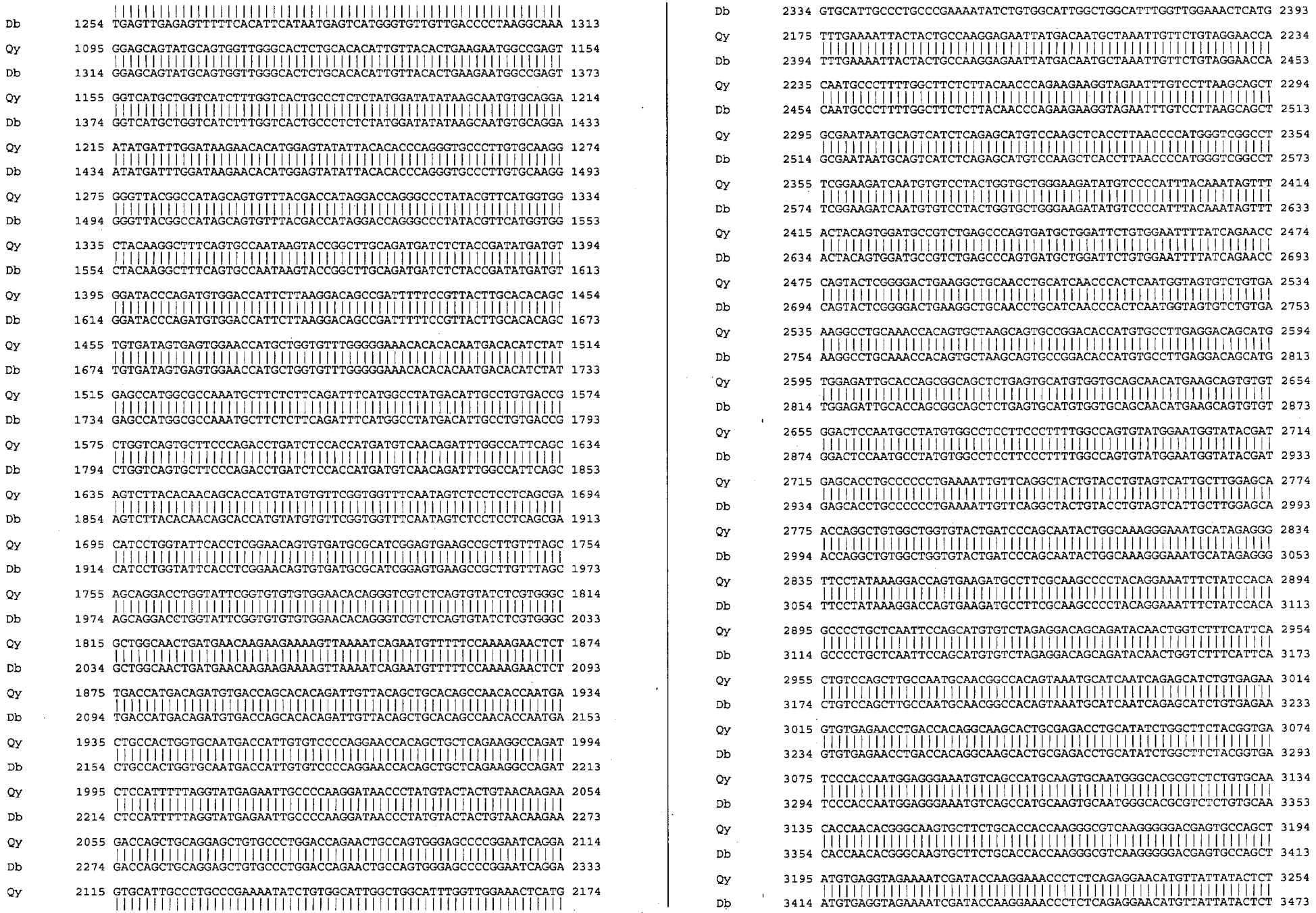
Qy       795 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 854
Db      1014 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 1073

Qy       855 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 914
Db      1074 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 1133

Qy       915 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 974
Db      1134 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 1193

Qy       975 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1034
Db      1194 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1253

Qy      1035 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1094
```


Db 1254 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1313
Qy 1095 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1154
Db 1314 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1373
Qy 1155 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1214
Db 1374 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1433
Qy 1215 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1274
Db 1434 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1493
Qy 1275 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1334
Db 1494 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1553
Qy 1335 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1394
Db 1554 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1613
Qy 1395 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1454
Db 1614 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1673
Qy 1455 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1514
Db 1674 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1733
Qy 1515 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1574
Db 1734 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1793
Qy 1575 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGC 1634
Db 1794 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGC 1853
Qy 1635 AGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGA 1694
Db 1854 AGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGA 1913
Qy 1695 CATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGC 1754
Db 1914 CATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGC 1973
Qy 1755 AGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGC 1814
Db 1974 AGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGC 2033
Qy 1815 GCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCT 1874
Db 2034 GCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCT 2093
Qy 1875 TGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGA 1934
Db 2094 TGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGA 2153
Qy 1935 CTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGAT 1994
Db 2154 CTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGAT 2213
Qy 1995 CTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAA 2054
Db 2214 CTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAA 2273
Qy 2055 GACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGA 2114
Db 2274 GACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGA 2333
Qy 2115 GTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATG 2174
Db 2334 GTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATG 2393
Qy 2175 TTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCA 2234
Db 2394 TTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCA 2453
Qy 2235 CAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCT 2294
Db 2454 CAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCT 2513
Qy 2295 GCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCT 2354
Db 2514 GCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCT 2573
Qy 2355 TCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTT 2414
Db 2574 TCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTT 2633
Qy 2415 ACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACC 2474
Db 2634 ACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACC 2693
Qy 2475 CAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGA 2534
Db 2694 CAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGA 2753
Qy 2535 AAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATG 2594
Db 2754 AAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATG 2813
Qy 2595 TGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGT 2654
Db 2814 TGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGT 2873
Qy 2655 GGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGAT 2714
Db 2874 GGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGAT 2933
Qy 2715 GAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCA 2774
Db 2934 GAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCA 2993
Qy 2775 ACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGG 2834
Db 2994 ACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGG 3053
Qy 2835 TTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACA 2894
Db 3054 TTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACA 3113
Qy 2895 GCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCA 2954
Db 3114 GCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCA 3173
Qy 2955 CTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAA 3014
Db 3174 CTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAA 3233
Qy 3015 GTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGA 3074
Db 3234 GTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGA 3293
Qy 3075 TCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAA 3134
Db 3294 TCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAA 3353
Qy 3135 CACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCT 3194
Db 3354 CACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCT 3413
Qy 3195 ATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCT 3254
Db 3414 ATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCT 3473

```
Qy       3255 TCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGC 3314
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3474 TCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGC 3533

Qy       3315 TATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGC 3374
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3534 TATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGC 3593

Qy       3375 CTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGC 3434
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3594 CTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGC 3653

Qy       3435 TGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTC 3494
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3654 TGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTC 3713

Qy       3495 TAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAA 3554
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3714 TAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAA 3773

Qy       3555 TTTCACCTGGCCCATCAAAATTCAGAT 3581
             |||||||||||||||||||||||||||
Db       3774 TTTCACCTGGCCCATCAAAATTCAGGT 3800
```

RESULT 5
AF106861
LOCUS AF106861 3819 bp mRNA linear PRI 06-JAN-1999
DEFINITION Homo sapiens attractin-2 (ATRN) mRNA, complete cds.
ACCESSION AF106861
VERSION AF106861.1 GI:4093195
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3819)
AUTHORS Duke-Cohan,J.S., Gu,J., Freeman,G.J. and Schlossman,S.F.
TITLE Cloning of cDNA for attractin-2, identical with that of attractin except for a GC-rich 222 bp 5' insertion
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3819)
AUTHORS Duke-Cohan,J.S., Gu,J., Freeman,G.J. and Schlossman,S.F.
TITLE Direct Submission
JOURNAL Submitted (15-NOV-1998) Division of Tumor Immunology, Dana Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, USA
FEATURES Location/Qualifiers
source 1. .3819
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
gene 1. .3819
/gene="ATRN"
CDS 1. .3819
/gene="ATRN"
/note="similar to Homo sapiens attractin in GenBank Accession Number AF034957"
/codon_start=1
/product="attractin-2"
/protein_id="AAD03057.1"
/db_xref="GI:4093196"
/translation="MVAAAAATEARLRRRTAATAALAGRSGGPHWDWDVTRAGRPGLG
AGLRLPRLLSPPLRPRLLLLLLLLPPPLLLLLLPCEAEAAAAAAAVSGSAAAEAKECD
RPCVNGGRCNPGTGQCVCPAGWVGEQCQHCGGRFRLTGSSGFVTDGPGNYKYKTKCTW
LIEGQPNRIMRLRFNHFATECSWDHLYVYDGDSIYAPLVAAFSGLIVPERDGNETVPE
VVATSGYALLHFFSDAAYNLTGFNITYSFDMCPNNCSGRGECKISNSSETVECECSEN
WKGEACDIPHCTDNCGFPHRGICNSSDVRGCSCFSDWQGPGCSVPVPANQSFWTREEY
SNLKLPRASHKAVVNGNIMWVVGGYMFNHSDYNMVLAYDLASREWLPLNRSVNNVVVR
YGHSLALYKDKIYMYGGKIDPTGNVTNELRVFHIHNESWVLLTPKAKEQYAVVGHSAH
IVTLKNGRVVMLVIFGHCPLYGYISNVQEYDLDKNTWSILHTQGALVQGGYGHSSVYD
HRTRALYVHGGYKAFSANKYRLADDLYRYDVDTQMWTILKDSRFFRYLHTAVIVSGTM
LVFGGNTHNDTSMSHGAKCFSSDFMAYDIACDRWSVLPRPDLHHDVNRFGHSAVLHNS
TMYVFGGFNSLLLSDILVFTSEQCDAHRSBAACLAAGPGIRCVWNTGSSQCISWALAT
DEQEEKLKSECFSKRTLDHDRCDQHTDCYSCTANTNDCHWCNDHCVPRNHSCSEGQIS
IFRYENCPKDNPMYYCNKKTSCRSCALDQNCQWEPRNQECIALPENICGIGWHLVGNS
CLKITTAKENYDNAKLFCRNHNALLASLTTQKKVEFVLKQLRIMQSSQSMSKLTLTPW
VGLRKINVSYWCWEDMSPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLN
GSVCERPANHSAKQCRTPCALRTACGDCTSGSSECMWCSNMKQCVDSNAYVASFPFGQ
CMEWYTMSTCPPENCSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKMPSQ
APTGNFYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCENLTTGKH
CETCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGDECQLCEVENRY
QGNPLRGTCYYTLLIDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASKNFN
LNITWAASFSAGTQAGEEMPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTW
PIKIQVQTEQ"

ORIGIN

Query Match 85.9%; Score 3492.4; DB 9; Length 3819;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 3508; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

```
Qy         48 GGCGGCGACGGCAGCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGTGTCAACGGCGGTCG 107
              | |||   | || ||   || || || |   | |  | || ||||||||||||||||||
Db        270 GTCGGGCTCAGCCGCAGCCGAGGCCAAGGAATGTGACCGGCCCTGTGTCAACGGCGGTCG 329

Qy        108 CTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCA 167
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        330 CTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCA 389

Qy        168 GCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGG 227
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        390 GCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGG 449

Qy        228 AAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAAT 287
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        450 AAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAAT 509

Qy        288 GAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGA 347
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        510 GAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGA 569

Qy        348 TGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAG 407
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        570 TGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAG 629

Qy        408 AGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTT 467
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        630 AGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTT 689

Qy        468 TTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTG 527
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        690 TTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTG 749

Qy        528 TCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGA 587
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        750 TCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGAAACTGTTGA 809

Qy        588 ATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA 647
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        810 ATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA 869

Qy        648 CTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTT 707
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        870 CTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTT 929

Qy        708 CTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGAC 767
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        930 CTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGAC 989

Qy        768 TCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGG 827
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        990 TCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGG 1049

Qy        828 AAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCT 887
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1050 AAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCT 1109
```

```
Qy       888 AGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGT 947
Db      1110 AGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGT 1169

Qy       948 TGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAA 1007
Db      1170 TGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAA 1229

Qy      1008 AATTGATTCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTC 1067
Db      1230 AATTGATCCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTC 1289

Qy      1068 ATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACA 1127
Db      1290 ATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACA 1349

Qy      1128 CATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCT 1187
Db      1350 CATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCT 1409

Qy      1188 CTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATT 1247
Db      1410 CTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATT 1469

Qy      1248 ACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAG 1307
Db      1470 ACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAG 1529

Qy      1308 GACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCT 1367
Db      1530 GACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCT 1589

Qy      1368 TGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAG 1427
Db      1590 TGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAG 1649

Qy      1428 CCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGG 1487
Db      1650 CCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGG 1709

Qy      1488 GGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTT 1547
Db      1710 GGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTT 1769

Qy      1548 CATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCA 1607
Db      1770 CATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCA 1829

Qy      1608 TGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGG 1667
Db      1830 TGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGG 1889

Qy      1668 TGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGC 1727
Db      1890 TGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGC 1949

Qy      1728 GCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACAC 1787
Db      1950 GCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACAC 2009

Qy      1788 AGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAA 1847
Db      2010 AGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAA 2069

Qy      1848 ATCAGAATGTTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTG 1907
Db      2070 ATCAGAATGTTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTG 2129

Qy      1908 TTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAG 1967
Db      2130 TTACAGCTGTACAGCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAG 2189

Qy      1968 GAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGA 2027
Db      2190 GAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGA 2249

Qy      2028 TAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAA 2087
Db      2250 TAACCCCATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAA 2309

Qy      2088 CTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCAT 2147
Db      2310 CTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCAT 2369

Qy      2148 TGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGA 2207
Db      2370 TGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGA 2429

Qy      2208 CAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAA 2267
Db      2430 CAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAA 2489

Qy      2268 GAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAA 2327
Db      2490 GAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAA 2549

Qy      2328 GCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGA 2387
Db      2550 GCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGA 2609

Qy      2388 AGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGC 2447
Db      2610 AGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGC 2669

Qy      2448 TGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCAT 2507
Db      2670 TGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCAT 2729

Qy      2508 CAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCG 2567
Db      2730 CAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCG 2789

Qy      2568 GACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCAT 2627
Db      2790 GACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCAT 2849

Qy      2628 GTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTT 2687
Db      2850 GTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTT 2909

Qy      2688 TGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTA 2747
Db      2910 TGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTA 2969

Qy      2748 CTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAA 2807
Db      2970 CTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAA 3029

Qy      2808 TACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCA 2867
Db      3030 TACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCA 3089

Qy      2868 AGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGA 2927
Db      3090 AGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGA 3149

Qy      2928 CAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAA 2987
Db      3150 CAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAA 3209

Qy      2988 ATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGA 3047
Db      3210 ATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGA 3269

Qy      3048 GACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAA 3107
```

```
Db       3270 GACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAA 3329

Qy       3108 GTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAA 3167
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3330 GTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAA 3389

Qy       3168 GGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCC 3227
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3390 GGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCC 3449

Qy       3228 TCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATC 3287
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3450 TCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATC 3509

Qy       3288 CCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAA 3347
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3510 CCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAA 3569

Qy       3348 CAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGC 3407
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3570 CAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGC 3629

Qy       3408 TGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAA 3467
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3630 TGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAA 3689

Qy       3468 CATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAA 3527
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3690 CATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAA 3749

Qy       3528 TATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGAT 3581
             ||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       3750 TATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGGT 3803
```

```
RESULT 6
AB062913
LOCUS       AB062913                4284 bp    mRNA    linear   ROD 16-JAN-2002
DEFINITION  Mesocricetus auratus Atrn mRNA for attractin, complete cds.
ACCESSION   AB062913
VERSION     AB062913.1  GI:16930100
KEYWORDS    .
SOURCE      Mesocricetus auratus (golden hamster)
  ORGANISM  Mesocricetus auratus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
            Mesocricetus.
REFERENCE   1
  AUTHORS   Kuramoto,T., Nomoto,T., Fujiwara,A., Mizutani,M., Sugimura,T. and
            Ushijima,T.
  TITLE     Insertional mutation of the Attractin gene in the black tremor
            hamster
  JOURNAL   Mamm. Genome 13 (1), 36-40 (2002)
  MEDLINE   21635551
   PUBMED   11773967
REFERENCE   2  (bases 1 to 4284)
  AUTHORS   Kuramoto,T. and Ushijima,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-JUN-2001) Takashi Kuramoto, National Cancer Center
            Research Institute, Carcinogenesis Division; Tsukiji 5-1-1,
            Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramot@gan2.ncc.go.jp,
            URL:http://www.ncc.go.jp/research/rat-genome/, Tel:81-3-3542-2511,
            Fax:81-3-5565-1753)
FEATURES             Location/Qualifiers
     source          1. .4284
                     /organism="Mesocricetus auratus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10036"
     gene            1. .4284
                     /gene="Atrn"
     CDS             1. .4284
                     /gene="Atrn"
                     /codon_start=1
                     /product="attractin"
                     /protein_id="BAB72012.1"
                     /db_xref="GI:16930101"
                     /translation="MVAAAAATEALLRGSTAATAAPAGRKVEQYRLCAATGAWRPGPR
                     AGLCLPWVVSRALPPPLLLLLLLFSLLLLPLPREAEAAAVSAAVSGSAAAEAKECDRP
                     CVNGGRCNPGTGQCVCPAGWVGEQCQHCGGRFRLTGSSGFVTDGPGNYKYKTCTWLI
                     EGQPNRIMRLRFNHFATECSWDHLYVYDGDSIYAPLIAAFSGLIVPERDGNETAPEVT
                     VTSGYALLHFFSDAAYNLTGFNITYNFDMCPNNCSGRGECKSSNSSNTVECECSENWK
                     GESCDIPHCTDNCGFPHRGICNASDVRGCSCSPHWQGPGCSIPVPANQSFWIREEYSD
                     FKLPRASHKAVVNGNIMWVVGGYMFNHSDYSMVLAYDLASREWLPLNHSVNNVVVRYG
                     HSLALHKDKIYMYGGKIDSTGNVTNELRVFHIHNESWVLLTPKAKEQYAVVGHSAHIV
                     TLATGRVVMLVIFGHCPLYGYISIVQEYDLEKNTWSILQTQGALVQGGYGHSSVYDHR
                     TKALYVHGGYKAFSANKYRLADDLYRYDVDTQMWTILKDSRFFRYLHTAVIVSGTMLV
                     FGGNTHNDTSMSHGAKCFSSDFMAYDIACDRWSVLPRPDLHHDVNRFGHSAVLHNSTM
                     YVFGGFNSLLLSDVLVFTSEQCDAHHSEAACVAAGPGIRCLWDTVSSRCTSWELATEE
                     QAEKLKSECFSKRTLDHDRCDQHTDCYSCTANTNDCHWCNDHCVPVNHSCTEGQISIS
                     KYDNCPKDNPMYYCNKKTSCRSCALDQNCQWEPRNQECIALPENICGIGWHLVGNSCL
                     KITTAKENYDNAKLSCRNHNAFLASLTSQKKVEFVLKHLRLMQSSQSMSKLTLTPWVG
                     LRKINVSYWCWEDMSPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLNGS
                     VCERPANHSAKQCRTPCALRTACSECTSSSSECMWCSNMKQCVDSNAYVASFPFGQCM
                     EWYTMSSCPPENCSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPIRMPSQAS
                     TGNVYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCENLTTGKHCE
                     TCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGDECQLCEVENRYQG
                     NPLKGTCYYTLLIDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASKNFNLN
                     ITWATSFSAGTQAGEEMPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTWPI
                     KIQIAFSQHSNFMDLVQFFVTFFSCFLSLLLVAAVVWKIKQSCWASRRREQLLREMQQ
                     MASRPFASVNVALETDEEPPDLIGGSIKTVPKPIALEPCFGNKAAVLSVFVRLPRGLG
                     GIPPPGQSGLAVASALVDISQQMPIVYKEKSGAVRNRKQQPPAQPGTCI"
ORIGIN

  Query Match           83.5%;  Score 3395.4;  DB 10;  Length 4284;
  Best Local Similarity 89.9%;  Pred. No. 0;
  Matches 3642;  Conservative    0;  Mismatches  411;  Indels    0;  Gaps    0;

Qy         16 GCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAGG 75
              | |||     |  ||   | ||  |  |  |  | | |   ||||  ||   || || || | |
Db        232 GAGGCCGAGGCCGCGGCGGTGTCGGCGGCGGTGTCAGGCTCGGCCGCAGCCGAGGCCAAG 291

Qy         76 AGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTC 135
                 | |  | ||  |||||||||||||| ||||||||||||||||||| ||||||||||||
Db        292 GAATGTGACCGGCCGTGTGTCAACGGCGGCCGCTGCAACCCTGGCACTGGCCAGTGCGTC 351

Qy        136 TGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGGCCGCTTCAGACTAACT 195
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 TGCCCCGCGGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGGCCGCTTCAGACTAACT 411

Qy        196 GGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCACG 255
              |||||||||| || || |||||||||||||| |||||||| || ||||||||||||||||
Db        412 GGATCTTCTGGATTCGTAACAGATGGACCTGGGAATTATAAGTATAAAACGAAGTGCACA 471

Qy        256 TGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACA 315
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 TGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACA 531

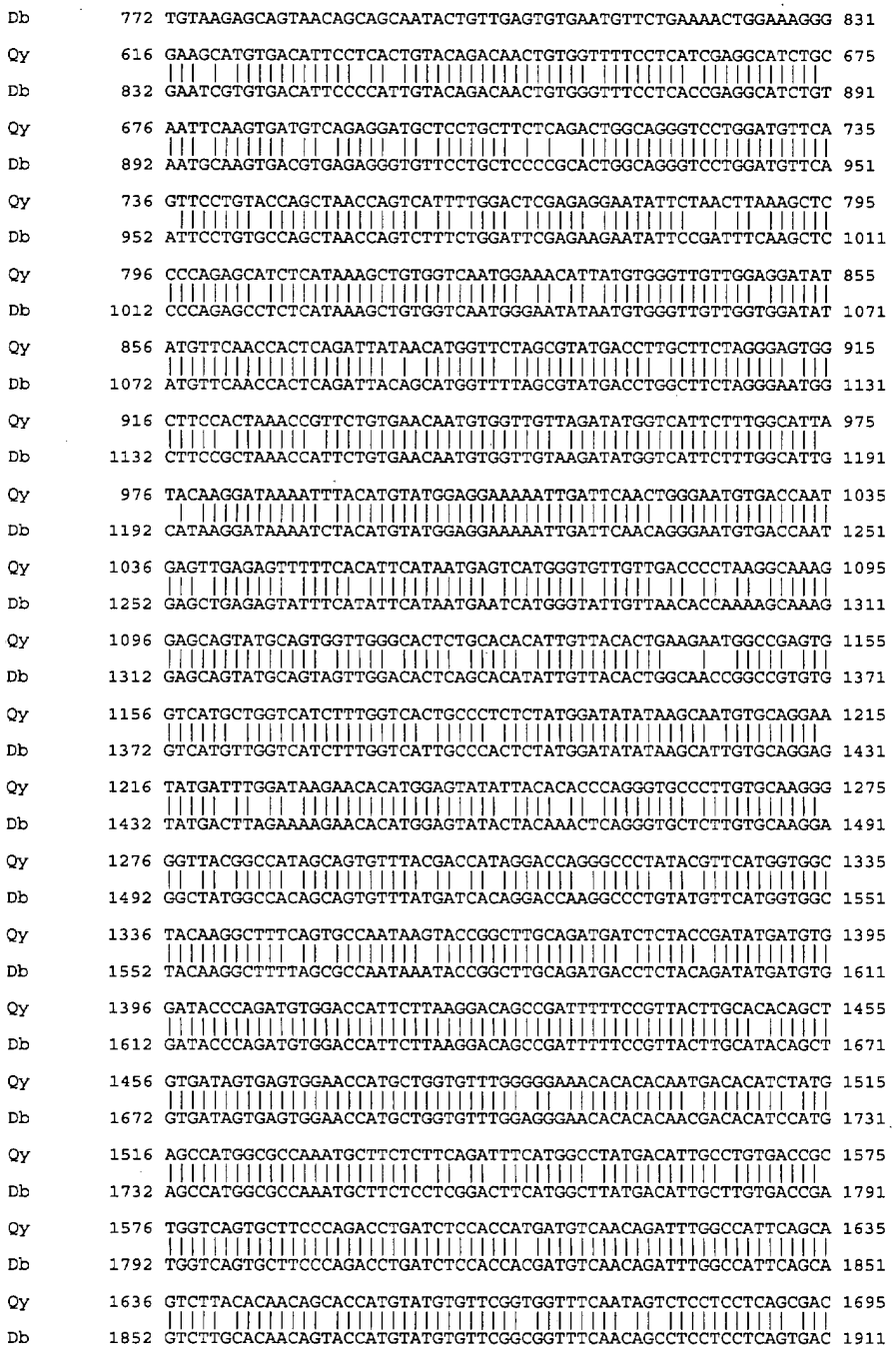
Qy        316 GAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTT 375
              || ||||| ||||||||||||||||||||||||||||||||||||||||| || ||
Db        532 GAATGTAGCTGGGACCATTTATATGTTTATGACGGGGACTCAATTTATGCACCTCTGATT 591

Qy        376 GCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTT 435
              ||||| |||||||| ||||||||||||||||||||||||||||||| | ||| |||||
Db        592 GCTGCCTTTAGTGGTCTCATTGTTCCTGAGAGAGATGGCAATGAGACAGCCCCAGAGGTC 651

Qy        436 GTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACT 495
              || ||| |||||||||||  |||||||||||| |||||||||||||| ||| |||||
Db        652 ACTGTCACTTCAGGTTATGCACTGCTGCATTTTTTCAGTGATGCTGCTTACAATCTGACT 711

Qy        496 GGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAG 555
              |||||||||| ||||||| |||||| |||||||| ||||| ||||||||||||||||||
Db        712 GGATTTAATATCACTTACAATTTTGACATGTGTCCGAATAATTGCTCAGGCCGAGGAGAA 771

Qy        556 TGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGT 615
              ||||||| |||||| |||||| |||||||||| ||||||||||||||||||||||||||
```


Db 772 TGTAAGAGCAGTAACAGCAGCAATACTGTTGAGTGTGAATGTTCTGAAAACTGGAAAGGG 831
Qy 616 GAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGC 675
Db 832 GAATCGTGTGACATTCCCCATTGTACAGACAACTGTGGGTTTCCTCACCGAGGCATCTGT 891
Qy 676 AATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCA 735
Db 892 AATGCAAGTGACGTGAGAGGGTGTTCCTGCTCCCCGCACTGGCAGGGTCCTGGATGTTCA 951
Qy 736 GTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTC 795
Db 952 ATTCCTGTGCCAGCTAACCAGTCTTTCTGGATTCGAGAAGAATATTCCGATTTCAAGCTC 1011
Qy 796 CCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATAT 855
Db 1012 CCCAGAGCCTCTCATAAAGCTGTGGTCAATGGGAATATAATGTGGGTTGTTGGTGGATAT 1071
Qy 856 ATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGG 915
Db 1072 ATGTTCAACCACTCAGATTACAGCATGGTTTTAGCGTATGACCTGGCTTCTAGGGAATGG 1131
Qy 916 CTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTA 975
Db 1132 CTTCCGCTAAACCATTCTGTGAACAATGTGGTTGTAAGATATGGTCATTCTTTGGCATTG 1191
Qy 976 TACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAAT 1035
Db 1192 CATAAGGATAAAATCTACATGTATGGAGGAAAAATTGATTCAACAGGGAATGTGACCAAT 1251
Qy 1036 GAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAG 1095
Db 1252 GAGCTGAGAGTATTTCATATTCATAATGAATCATGGGTATTGTTAACACCAAAAGCAAAG 1311
Qy 1096 GAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTG 1155
Db 1312 GAGCAGTATGCAGTAGTTGGACACTCAGCACATATTGTTACACTGGCAACCGGCCGTGTG 1371
Qy 1156 GTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAA 1215
Db 1372 GTCATGTTGGTCATCTTTGGTCATTGCCCACTCTATGGATATATAAGCATTGTGCAGGAG 1431
Qy 1216 TATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGG 1275
Db 1432 TATGACTTAGAAAAGAACACATGGAGTATACTACAAACTCAGGGTGCTCTTGTGCAAGGA 1491
Qy 1276 GGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGC 1335
Db 1492 GGCTATGGCCACAGCAGTGTTTATGATCACAGGACCAAGGCCCTGTATGTTCATGGTGGC 1551
Qy 1336 TACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTG 1395
Db 1552 TACAAGGCTTTTAGCGCCAATAAATACCGGCTTGCAGATGACCTCTACAGATATGATGTG 1611
Qy 1396 GATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCT 1455
Db 1612 GATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCATACAGCT 1671
Qy 1456 GTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATG 1515
Db 1672 GTGATAGTGAGTGGAACCATGCTGGTGTTTGGAGGGAACACACACAACGACACATCCATG 1731
Qy 1516 AGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGC 1575
Db 1732 AGCCATGGCGCCAAATGCTTCTCCTCGGACTTCATGGCTTATGACATTGCTTGTGACCGA 1791
Qy 1576 TGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCA 1635
Db 1792 TGGTCAGTGCTTCCCAGACCTGATCTCCACCACGATGTCAACAGATTTGGCCATTCAGCA 1851
Qy 1636 GTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGAC 1695
Db 1852 GTCTTGCACAACAGTACCATGTATGTGTTCGGCGGTTTCAACAGCCTCCTCCTCAGTGAC 1911
Qy 1696 ATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCA 1755
Db 1912 GTCTTAGTCTTTACCTCGGAGCAATGTGACGCCCACCACAGTGAGGCAGCTTGTGTAGCA 1971
Qy 1756 GCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCG 1815
Db 1972 GCAGGACCTGGTATCCGGTGTCTGTGGGACACAGTGTCATCTCGATGTACCTCCTGGGAG 2031
Qy 1816 CTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTT 1875
Db 2032 TTGGCAACTGAAGAACAAGCAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACCCTT 2091
Qy 1876 GACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGAC 1935
Db 2092 GACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAATACCAATGAC 2151
Qy 1936 TGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATC 1995
Db 2152 TGCCACTGGTGCAATGATCACTGTGTCCCTGTGAACCACAGCTGCACAGAAGGCCAGATC 2211
Qy 1996 TCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAAG 2055
Db 2212 TCCATTTCCAAGTATGACAATTGCCCCAAGGATAACCCCATGTACTACTGCAATAAGAAA 2271
Qy 2056 ACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAG 2115
Db 2272 ACAAGCTGCAGAAGCTGTGCCTTAGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAG 2331
Qy 2116 TGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGT 2175
Db 2332 TGCATCGCTCTGCCTGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGT 2391
Qy 2176 TTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCAC 2235
Db 2392 CTGAAAATCACTACTGCTAAGGAAAATTATGACAATGCTAAGTTGTCCTGTAGAAACCAC 2451
Qy 2236 AATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTG 2295
Db 2452 AATGCCTTTTTGGCTTCCCTTACATCCCAGAAGAAGGTGGAATTCGTCCTTAAGCATCTT 2511
Qy 2296 CGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTT 2355
Db 2512 CGATTAATGCAGTCATCTCAGAGCATGTCCAAGCTCACTCTGACTCCATGGGTTGGTCTT 2571
Qy 2356 CGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTA 2415
Db 2572 CGGAAGATCAATGTGTCTTATTGGTGCTGGGAGGATATGTCTCCATTTACAAATAGTTTG 2631
Qy 2416 CTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCC 2475
Db 2632 CTGCAGTGGATGCCATCCGAGCCCAGTGATGCTGGTTTCTGTGGGATCTTGTCAGAGCCT 2691
Qy 2476 AGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAA 2535
Db 2692 AGTACTCGGGGATTGAAGGCTGCAACCTGCATCAACCCACTCAATGGCAGTGTTTGTGAA 2751
Qy 2536 AGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGT 2595
Db 2752 AGGCCTGCAAACCACAGTGCCAAGCAGTGCCGGACACCGTGTGCCCTGCGGACAGCGTGT 2811
Qy 2596 GGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTG 2655
Db 2812 AGTGAGTGCACTAGCAGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTA 2871
Qy 2656 GACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATG 2715
Db 2872 GACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAGTGGTATACAATG 2931
Qy 2716 AGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAA 2775
Db 2932 AGCAGCTGCCCACCTGAAAATTGCTCTGGCTACTGTACCTGCAGCCATTGCTTGGAGCAG 2991

```
Qy       2776 CCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGT 2835
Db       2992 CCAGGCTGTGGCTGGTGTACTGATCCTAGCAATACTGGGAAAGGGAAATGTATCGAGGGC 3051

Qy       2836 TCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAG 2895
Db       3052 AGCTATAAAGGACCCATAAGGATGCCTTCACAGGCTTCTACAGGAAATGTGTATCCACAG 3111

Qy       2896 CCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCAC 2955
Db       3112 CCTCTTCTGAACTCTAGCATGTGTTTAGAGGACAGCAGATACAACTGGTCTTTCATTCAC 3171

Qy       2956 TGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAG 3015
Db       3172 TGTCCAGCTTGCCAGTGCAATGGACACAGCAAATGTATTAACCAGAGTATCTGTGAGAAG 3231

Qy       3016 TGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGAT 3075
Db       3232 TGTGAGAACCTGACCACTGGCAAGCACTGTGAGACCTGCATATCTGGCTTCTATGGTGAT 3291

Qy       3076 CCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAAC 3135
Db       3292 CCGACTAATGGAGGCAAATGTCAGCCATGTAAGTGCAACGGGCATGCATCGCTGTGCAAT 3351

Qy       3136 ACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTA 3195
Db       3352 ACCAACACCGGCAAGTGCTTCTGCACCACCAAGGGTGTCAAGGGAGATGAGTGCCAGCTA 3411

Qy       3196 TGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTT 3255
Db       3412 TGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAAAGGAACATGTTATTATACCCTT 3471

Qy       3256 CTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCT 3315
Db       3472 CTCATTGACTATCAGTTCACTTTTAGCCTGTCCCAGGAAGACGATCGCTATTATACAGCC 3531

Qy       3316 ATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCC 3375
Db       3532 ATCAATTTTGTGGCCACTCCTGATGAACAAAACAGGGATTTGGACATGTTTATCAATGCC 3591

Qy       3376 TCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCT 3435
Db       3592 TCAAAAAACTTCAACCTCAACATCACCTGGGCTACCAGCTTCTCAGCTGGAACCCAGGCA 3651

Qy       3436 GGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCT 3495
Db       3652 GGAGAAGAAATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGCTTCTCT 3711

Qy       3496 AATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAAT 3555
Db       3712 AATGAGAAATTTGATTTTCGCAACCATCCAAACATTACTTTCTTTGTTTATGTCAGTAAT 3771

Qy       3556 TTCACCTGGCCCATCAAAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCTG 3615
Db       3772 TTCACCTGGCCCATCAAAATTCAGATTGCCTTCTCCCAGCACAGCAATTTTATGGACCTG 3831

Qy       3616 GTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGTG 3675
Db       3832 GTACAGTTCTTCGTGACTTTCTTCAGCTGTTTCCTCTCTCTGCTCCTAGTGGCTGCAGTG 3891

Qy       3676 GTTTGGAAGATCAAACAAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGAG 3735
Db       3892 GTTTGGAAGATCAAGCAGAGCTGCTGGGCATCCAGGAGGAGGGAGCAACTTCTTCGGGAG 3951

Qy       3736 ATGCAACAGATGGCCAGCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGAG 3795
Db       3952 ATGCAGCAGATGGCCAGCCGTCCCTTTGCTTCTGTAAATGTTGCCTTGGAAACAGATGAA 4011

Qy       3796 GAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGAG 3855
Db       4012 GAGCCTCCTGATCTTATTGGGGGAAGTATAAAGACTGTCCCCAAGCCCATTGCCCTGGAG 4071

Qy       3856 CCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCTG 3915

Db       4072 CCCTGCTTTGGTAACAAAGCTGCAGTCCTCTCTGTATTTGTGAGGCTCCCTCGAGGACTG 4131

Qy       3916 GGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACATT 3975
Db       4132 GGTGGAATCCCTCCTCCTGGTCAGTCAGGTCTTGCTGTGGCCAGTGCCCTGGTGGACATT 4191

Qy       3976 TCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCAG 4035
Db       4192 TCTCAACAGATGCCGATAGTATACAAAGAGAAGTCAGGAGCGGTAAGAAACCGGAAGCAG 4251

Qy       4036 CAGCCCCCTGCACAGCCTGGGACCTGCATCTGA 4068
Db       4252 CAGCCCCCTGCACAGCCTGGAACCTGCATTTGA 4284
```

```
RESULT 7
AF116897
LOCUS       AF116897               6083 bp    mRNA    linear   ROD 20-MAR-1999
DEFINITION  Mus musculus mahogany protein mRNA, complete cds.
ACCESSION   AF116897
VERSION     AF116897.1  GI:4454560
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 6083)
  AUTHORS   Nagle,D.L., McGrail,S.H., Vitale,J., Woolf,E.A., Dussault,B.J. Jr.,
            DiRocco,L., Holmgren,L., Montagno,J., Bork,P., Huszar,D.,
            Fairchild-Huntress,V., Ge,P., Keilty,J., Ebeling,C., Baldini,L.,
            Gilchrist,J., Burn,P., Carlson,G.A. and Moore,K.J.
  TITLE     The mahogany protein is a receptor involved in suppression of
            obesity
  JOURNAL   Nature 398 (6723), 148-152 (1999)
  MEDLINE   99184159
   PUBMED   10086355
REFERENCE   2  (bases 1 to 6083)
  AUTHORS   Nagle,D.L., McGrail,S.H., Vitale,J., Woolf,E.A., Dussault,B.J. Jr.,
            DiRocco,L., Holmgren,L., Montagno,J., Bork,P., Huszar,D.,
            Fairchild-Huntress,V., Ge,P., Keilty,J., Ebeling,C., Baldini,L.,
            Gilchrist,J., Burn,P., Carlson,G.A. and Moore,K.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-DEC-1998) Sequence Analysis, Millennium
            Pharmaceuticals, 640 Memorial Drive, Cambridge, MA 02139, USA
FEATURES             Location/Qualifiers
     source          1. .6083
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
     CDS             70. .4488
                     /note="MG"
                     /codon_start=1
                     /product="mahogany protein"
                     /protein_id="AAD20947.1"
                     /db_xref="GI:4454561"
                     /translation="MVAVAAAAATEARLRGSTTTTAAPAGRKGRQHRPCTATGAWRPG
                     PRARLCLPRVLSRALPPPPLLPLLFSLLLLPLPREAEAAAVAAAVSGSAAAEAKECDR
                     PCVNGGRCNPGTGQCVCPTGWVGEQCQHCGGRFRLTGSSGFVTDGPGNYKYKTKCTWL
                     IEGQPNRIMRLRFNHFATECSWDHLYVYDGDSIYAPLIAAFSGLIVPERDGNETAPEV
                     TVTSGYALLHFFSDAAYNLTGFNITYNFDMCPNNCSGRGECKSSNSSSAVECECSENW
                     KGESCDIPHCTDNCGFPHRGICNASDTRGCSCFPHWQGPGCSIPVPANQSFWTREEYS
                     DLKLPRASHKAVVNGNIMWVVGGYMFNHSDYSMVLAYDLTSREWLPLNHSVNSVVVRY
                     GHSLALHKDKIYMYGGKIDSTGNVTNELRVFHIHNESWVLLTPKAKDQYAVVGHSAHI
                     VTLASGRVVMLVIFGHCPLYGYISVVQEYDLEKNTWSILHTQGALVQGGYGHSSVYDD
                     RTKALYVHGGYKAFSANKYRLADDLYRYDVDTQMWTILKDSRFFRYLHTAVIVSGTML
                     VFGGNTHNDTSMSHGAKCFSSDFMAYDIACDRWSVLPRPELHHDVNRFGHSAVLYNST
                     MYVFGGFNSLLLSDVLVFTSEQCDAHRSEAACVAAGPGIRCLWDTQSSRCTSWELATE
                     EQAEKLKSECFSKRTLDHDRCDQHTDCYSCTANTNDCHWCNDHCVPVNHSCTEGQISI
                     AKYESCPKDNPMYYCNKKTSCRSCALDQNCQWEPRNQECIALPENICGNGWHLVGNSC
                     LKITTAKENYDNAKLSCRNHNAFLASLTSQKKVEFVLKQLRLMQSSQSMSKLTLTPWV
                     GLRKINVSYWCWEDMSPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLNG
                     SVCERPANHSAKQCRTPCALRTACGECTSSSSECMWCSNMKQCVDSNAYVASFPFGQC
```

MEWYTMSSCPPENCSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKMPSQA
SAGNVYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCEDLTTGKHC
ETCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGDECQLCEVENRYQ
GNPLKGTCYYTLLIDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASKNFNL
NITWATSFPAGTQTGEEVPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTWP
IKIQIAFSQHSNFMDLVQFFVTFFSCFLSLLLVAAVVWKIKQSCWASRRREQLLREMQ
QMASRPFASVNVALETDEEPPDLIGGSIKTVPKPIALEPCFGNKAAVLSVFVRLPRGL
GGIPPPGQSRSRCGQCPGGHFSADANSVQGEVRSCKKPEAAAACTAWNLHLILGQEFA
LTEGVWSTKLSTGEAAGRKQAWGRRLGTLEASNSCA"

ORIGIN

Query Match 83.0%; Score 3377.4; DB 10; Length 6083;
Best Local Similarity 89.7%; Pred. No. 0;
Matches 3638; Conservative 0; Mismatches 416; Indels 1; Gaps 1;

```
Qy        15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74
Db       303 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAA 362

Qy        75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134
Db       363 GGAATGTGACCGGCCGTGTGTCAACGGCGGCCGCTGCAACCCTGGCACCGGCCAGTGCGT 422

Qy       135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194
Db       423 CTGCCCCACGGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 482

Qy       195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254
Db       483 TGGCTCTTCTGGATTTGTAACAGATGGACCTGGGAATTATAAATATAAGACGAAGTGCAC 542

Qy       255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314
Db       543 ATGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGCTTCAACCATTTTGCTAC 602

Qy       315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374
Db       603 AGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATCTACGCACCTCTGAT 662

Qy       375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434
Db       663 TGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCTCCTGAGGT 722

Qy       435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494
Db       723 CACTGTCACTTCAGGTTATGCACTGCTGCATTTTTTCAGTGATGCTGCTTATAATCTGAC 782

Qy       495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554
Db       783 TGGATTTAATATCACTTACAATTTTGACATGTGTCCGAATAATTGCTCAGGCCGAGGAGA 842

Qy       555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614
Db       843 GTGTAAGAGCAGTAACAGCAGCAGCGCTGTTGAGTGTGAATGTTCTGAAAACTGGAAAGG 902

Qy       615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 674
Db       903 GGAGTCGTGTGACATTCCTCACTGTACAGACAACTGTGGCTTTCCTCACCGAGGCATCTG 962

Qy       675 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 734
Db       963 TAATGCAAGCGATACCAGAGGGTGCTCCTGCTTTCCTCACTGGCAGGGTCCTGGATGTTC 1022

Qy       735 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 794
Db      1023 AATTCCTGTGCCAGCTAACCAGTCTTTTTGGACTCGAGAAGAATATTCTGATTTAAAGCT 1082

Qy       795 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 854
Db      1083 TCCCAGAGCCTCTCATAAAGCTGTGGTCAATGGAAATATAATGTGGGTTGTTGGCGGATA 1142

Qy       855 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 914
Db      1143 TATGTTCAACCATTCAGATTACAGCATGGTTTTAGCGTATGACCTGACTTCTAGGGAATG 1202

Qy       915 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 974
Db      1203 GCTTCCACTAAACCATTCTGTGAACAGTGTGGTTGTAAGATATGGTCATTCTTTGGCATT 1262

Qy       975 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1034
Db      1263 ACATAAGGATAAAATCTACATGTATGGAGGAAAAATTGATTCAACAGGGAACGTGACCAA 1322

Qy      1035 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1094
Db      1323 TGAGCTGAGAGTATTTCATATTCATAATGAATCATGGGTATTGTTAACTCCGAAAGCTAA 1382

Qy      1095 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1154
Db      1383 GGATCAGTATGCAGTGGTTGGACACTCAGCACACATTGTTACACTGGCATCTGGCCGTGT 1442

Qy      1155 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1214
Db      1443 GGTCATGTTGGTCATCTTCGGTCATTGCCCACTCTATGGATATATAAGCGTTGTGCAGGA 1502

Qy      1215 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1274
Db      1503 ATATGACTTGGAAAAGAACACATGGAGTATATTACATACTCAGGGTGCTCTTGTGCAAGG 1562

Qy      1275 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1334
Db      1563 GGGTTATGGCCACAGTAGTGTTTATGATGACAGGACCAAGGCTCTGTACGTTCATGGTGG 1622

Qy      1335 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1394
Db      1623 CTACAAGGCTTTCAGCGCCAACAAATACCGGCTTGCAGATGACCTCTACAGATACGATGT 1682

Qy      1395 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1454
Db      1683 GGATACTCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCATACAGC 1742

Qy      1455 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1514
Db      1743 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGAGGGAACACACACAATGACACTTCCAT 1802

Qy      1515 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1574
Db      1803 GAGCCACGGTGCCAAATGCTTCTCCTCGGACTTCATGGCTTATGACATTGCTTGTGACCG 1862

Qy      1575 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGC 1634
Db      1863 ATGGTCAGTGCTTCCCAGACCTGAGCTCCATCATGATGTCAACAGATTTGGCCATTCAGC 1922

Qy      1635 AGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGA 1694
Db      1923 AGTCTTGTACAACAGCACCATGTATGTGTTCGGCGGCTTCAACAGCCTCCTCCTCAGTGA 1982

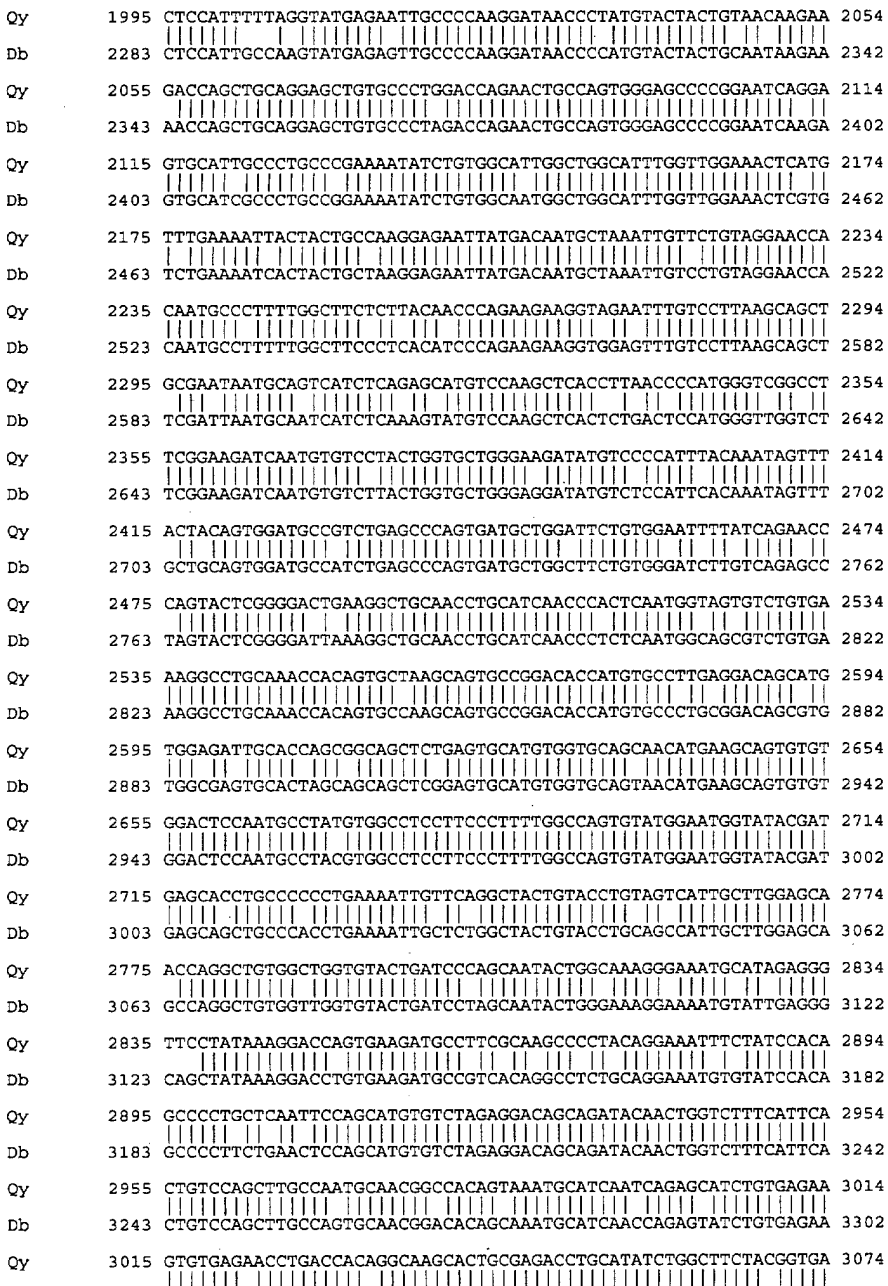
Qy      1695 CATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGC 1754
Db      1983 CGTCTTGGTCTTTACCTCGGAGCAGTGCGATGCACACCGCAGTGAAGCTGCTTGTGTGGC 2042

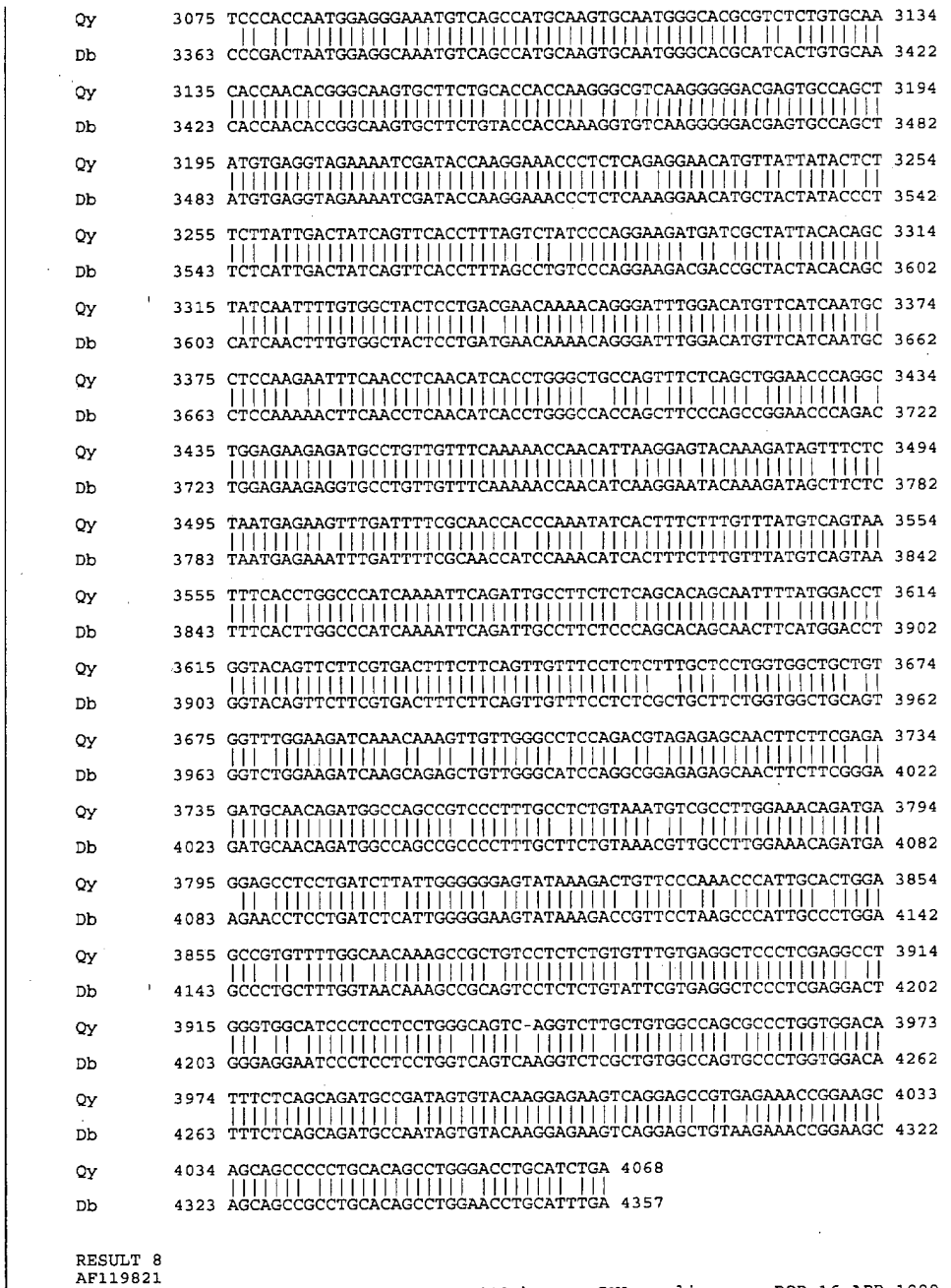
Qy      1755 AGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGC 1814
Db      2043 AGCAGGACCTGGTATCCGGTGTCTGTGGGACACACAGTCGTCTCGATGTACCTCCTGGGA 2102

Qy      1815 GCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCT 1874
Db      2103 GTTGGCAACTGAAGAACAAGCAGAAAAGTTAAAATCAGAGTGTTTTTCTAAAAGAACCCT 2162

Qy      1875 TGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGA 1934
Db      2163 TGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAATACCAATGA 2222

Qy      1935 CTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGAT 1994
Db      2223 CTGCCACTGGTGCAATGATCACTGTGTCCCTGTGAACCACAGCTGCACAGAAGGCCAGAT 2282
```





RESULT 8
AF119821
LOCUS AF119821 4313 bp mRNA linear ROD 16-APR-1999
DEFINITION Mus musculus attractin (Mgca) mRNA, complete cds.

```
ACCESSION   AF119821
VERSION     AF119821.1  GI:4585306
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 4313)
  AUTHORS   Gunn,T.M., Miller,K.A., He,L., Hyman,R.W., Davis,R.W., Azarani,A.,
            Schlossman,S.F., Duke-Cohan,J.S. and Barsh,G.S.
  TITLE     The mouse mahogany locus encodes a transmembrane form of human
            attractin
  JOURNAL   Nature 398 (6723), 152-156 (1999)
  MEDLINE   99184160
   PUBMED   10086356
REFERENCE   2  (bases 1 to 4313)
  AUTHORS   Gunn,T.M., Miller,K.A., He,L., Hyman,R.W., Davis,R.W., Azarani,A.,
            Schlossman,S.F., Duke-Cohan,J.S. and Barsh,G.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JAN-1999) HHMI, Stanford, Beckman Center B271A,
            Stanford, CA 94305-5323, USA
FEATURES             Location/Qualifiers
     source          1. .4313
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /chromosome="2"
     gene            1. .4313
                     /gene="Mgca"
                     /note="mahogany candidate"
     CDS             27. .4313
                     /gene="Mgca"
                     /note="mahogany"
                     /codon_start=1
                     /product="attractin"
                     /protein_id="AAD25372.1"
                     /db_xref="GI:4585307"
                     /translation="MVAVAVAASTEARLRGSTTATAAPAGRKGRQHRPCTATGAWRPG
                     PRARLCLPRVLSRALPPPPLLPLLFSLLLLPLPREAEAAAVAAAVSGSAAAEAKECDR
                     PCVNGGRCNPGTGQCVCPTGWVGEQCQHCGGRFRLTGSSGFVTDGPGNYKYKTKCTWL
                     IEGYPNAVLRLRFNHFATECSWDHLYVYDGDSIYAPLIAAFSGLIVPERDGNETAPEV
                     TVTSGYALLHFFSDAAYNLTGFNITYNFDMCPNNCSARGECKSSNSSSAVECECSENW
                     KGESCDIPHCTDNCGFPHRGICNASDTRGCSCFPHWQGPGCSIPVPANQSFWTREEYS
                     DLKLPRASHKAVVNGNIMWVVGGYMFNHSDYSMVLAYDLTSREWLPLNHSVNSVVVRY
                     GHSLALHKDKIYMYGGKIDSTGNVTNELRVFHIHNESWVLLTPKAKDQYAVVGHSAHI
                     VTLASGRVVVMLVIFGHCPLYGYISVVQEYDLEKNTWSILHTQGALVQGGYGHSSAYDD
                     RTKALYVHGGYKAFSANKYRLADDLYRYDVDTQMWTILKDSRFFRYLHTAVIVSGTML
                     VFGGNTHNDTSMSHGAKCFSSDFMAYDIACDRWSVLPRPELHHDVNRFGHSAVLYNST
                     MYVFGGFNSLLLSDVLVFTSEQCDAHRSEAACVAAGPGIRCLWDTQSSRCTSWELATE
                     EQAEKLKSECFSKRTLDHDRCDQHTDCYSCTANTNDCHWCNDHCVPVNHSCTEGQISI
                     AKYESCPKDNPMYYCNKKTSCRSCALDQNCQWEPRNQECIALPENICGNGWHLVGNSC
                     LKITTAKENYDNAKLSCRNHNAFLASLTSQKKVELVLKQLRLMQSSQSMSKLTLTPWV
                     GLRKINVSYWCWEDMSPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLNG
                     SVCERPANHSAKQCRTPCALRTACGECTSSSSECMWCSNMKQCVDSNAYVASFPFGQC
                     MEWYTMSSCPPENCSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKMPSQA
                     SAGNVYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSVCEKCEDLTTGKHC
                     ETCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGDECQLCKVENRYQ
                     GNPLKGTCYYTLLIDYQFTFSLSQGDDRYYTAINFVATPDEQNRDFDMFINASKKFNL
                     NITWATSFPAGTQTGEEVPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTWP
                     IKIQIAFSQHSNFMDLVQFFVTFFSCFLSLLLVAAVVWKIKQSCWASRRREQLLREMQ
                     QMASRPFASVNVALETDEEPPDLIGGSIKTVPKPIALEPCFGNKAAVLSVFVRLPRGL
                     GGIPPPGQSGLAVASALVDISQQMPIVYKEKSGAVRNRKQQPPAQPGTCI"
ORIGIN

  Query Match           82.5%;  Score 3354.8;  DB 10;  Length 4313;
  Best Local Similarity  89.2%;  Pred. No. 0;
  Matches 3617;  Conservative    0;  Mismatches  437;  Indels    0;  Gaps    0;

Qy         15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74
Db        260 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCGGGCTCGGCCGCAGCCGAGGCCAA 319

Qy         75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134
Db        320 GGAATGTGACCGGCCGTGTGTCAACGGCGGCCGCTGCAACCCTGGCACCGGCCAGTGCGT 379

Qy        135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194
Db        380 CTGCCCCACGGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 439

Qy        195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254
Db        440 TGGCTCTTCTGGATTTGTAACAGATGGACCTGGGAATTATAAATATAAGACGAAGTGCAC 499

Qy        255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314
Db        500 ATGGCTCATTGAAGGCTATCCAAATGCAGTGCTAAGGTTAAGATTCAACCATTTTGCTAC 559

Qy        315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374
Db        560 AGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATCTACGCACCTCTGAT 619

Qy        375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434
Db        620 TGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCTCCTGAGGT 679

Qy        435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494
Db        680 CACTGTCACTTCAGGTTATGCACTGCTGCATTTTTTCAGTGATGCTGCTTATAATCTGAC 739

Qy        495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554
Db        740 TGGATTTAATATCACTTACAATTTTGACATGTGTCCGAATAATTGCTCAGCCCGAGGAGA 799

Qy        555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614
Db        800 GTGTAAGAGCAGTAACAGCAGCAGCGCTGTTGAGTGTGAATGTTCTGAAAACTGGAAAGG 859

Qy        615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 674
Db        860 GGAGTCGTGTGACATTCCTCACTGTACAGACAACTGTGGCTTTCCTCACCGAGGCATCTG 919

Qy        675 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 734
Db        920 TAATGCAAGCGATACCAGAGGGTGCTCCTGCTTTCCTCACTGGCAGGGTCCTGGATGCTC 979

Qy        735 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 794
Db        980 AATTCCTGTGCCAGCTAACCAGTCTTTTTGGACTCGAGAAGAATATTCTGATTTAAAGCT 1039

Qy        795 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 854
Db       1040 TCCCAGAGCCTCTCATAAAGCTGTGGTCAATGGAAATATAATGTGGGTTGTTGGCGGATA 1099

Qy        855 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 914
Db       1100 TATGTTCAACCATTCAGATTACAGCATGGTTTTAGCGTATGACCTGACTTCTAGGGAATG 1159

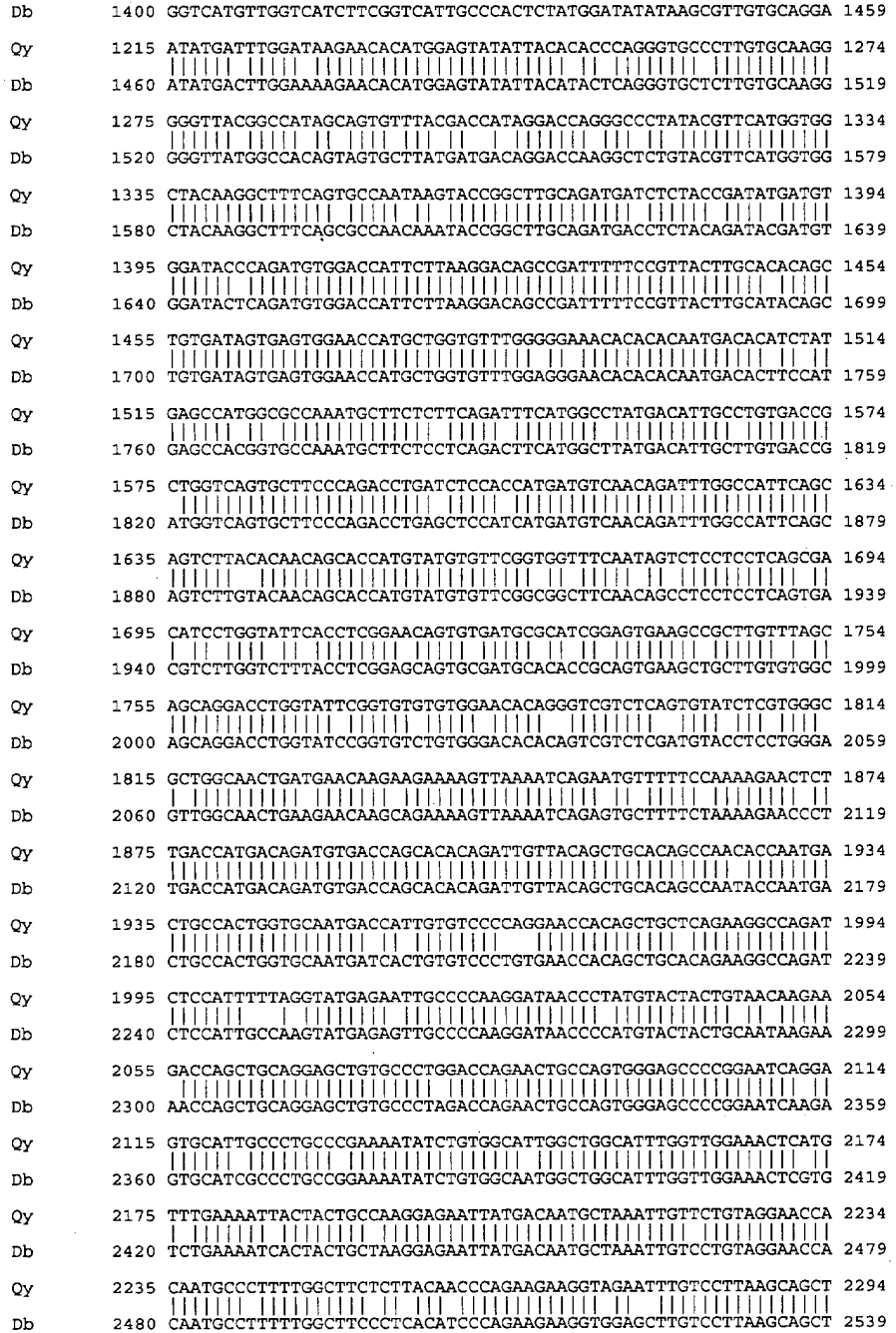
Qy        915 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 974
Db       1160 GCTTCCACTAAACCATTCTGTGAACAGTGTGGTTGTAAGATATGGTCATTCTTTGGCATT 1219

Qy        975 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1034
Db       1220 ACATAAGGATAAAATCTACATGTATGGAGGAAAAATTGATTCAACAGGGAACGTGACCAA 1279

Qy       1035 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1094
Db       1280 TGAGCTGAGAGTATTTCATATTCATAATGAATCATGGGTATTGTTAACTCCGAAAGCTAA 1339

Qy       1095 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1154
Db       1340 GGATCAGTATGCAGTGGTTGGACACTCAGCACACATTGTTACACTGGCATCTGGCCGTGT 1399

Qy       1155 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1214
```


Db 1400 GGTCATGTTGGTCATCTTCGGTCATTGCCCACTCTATGGATATATAAGCGTTGTGCAGGA 1459
Qy 1215 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1274
Db 1460 ATATGACTTGGAAAAGAACACATGGAGTATATTACATACTCAGGGTGCTCTTGTGCAAGG 1519
Qy 1275 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1334
Db 1520 GGGTTATGGCCACAGTAGTGCTTATGATGACAGGACCAAGGCTCTGTACGTTCATGGTGG 1579
Qy 1335 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1394
Db 1580 CTACAAGGCTTTCAGCGCCAACAAATACCGGCTTGCAGATGACCTCTACAGATACGATGT 1639
Qy 1395 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1454
Db 1640 GGATACTCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCATACAGC 1699
Qy 1455 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1514
Db 1700 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGAGGGAACACACACAATGACACTTCCAT 1759
Qy 1515 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1574
Db 1760 GAGCCACGGTGCCAAATGCTTCTCCTCAGACTTCATGGCTTATGACATTGCTTGTGACCG 1819
Qy 1575 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGC 1634
Db 1820 ATGGTCAGTGCTTCCCAGACCTGAGCTCCATCATGATGTCAACAGATTTGGCCATTCAGC 1879
Qy 1635 AGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGA 1694
Db 1880 AGTCTTGTACAACAGCACCATGTATGTGTTCGGCGGCTTCAACAGCCTCCTCCTCAGTGA 1939
Qy 1695 CATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGC 1754
Db 1940 CGTCTTGGTCTTTACCTCGGAGCAGTGCGATGCACACCGCAGTGAAGCTGCTTGTGTGGC 1999
Qy 1755 AGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGC 1814
Db 2000 AGCAGGACCTGGTATCCGGTGTCTGTGGGACACACAGTCGTCTCGATGTACCTCCTGGGA 2059
Qy 1815 GCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCT 1874
Db 2060 GTTGGCAACTGAAGAACAAGCAGAAAAGTTAAAATCAGAGTGCTTTTCTAAAAGAACCCT 2119
Qy 1875 TGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGA 1934
Db 2120 TGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAATACCAATGA 2179
Qy 1935 CTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGAT 1994
Db 2180 CTGCCACTGGTGCAATGATCACTGTGTCCCTGTGAACCACAGCTGCACAGAAGGCCAGAT 2239
Qy 1995 CTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAA 2054
Db 2240 CTCCATTGCCAAGTATGAGAGTTGCCCCAAGGATAACCCCATGTACTACTGCAATAAGAA 2299
Qy 2055 GACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGA 2114
Db 2300 AACCAGCTGCAGGAGCTGTGCCCTAGACCAGAACTGCCAGTGGGAGCCCCGGAATCAAGA 2359
Qy 2115 GTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATG 2174
Db 2360 GTGCATCGCCCTGCCGGAAAATATCTGTGGCAATGGCTGGCATTTGGTTGGAAACTCGTG 2419
Qy 2175 TTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCA 2234
Db 2420 TCTGAAAATCACTACTGCTAAGGAGAATTATGACAATGCTAAATTGTCCTGTAGGAACCA 2479
Qy 2235 CAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCT 2294
Db 2480 CAATGCCTTTTTGGCTTCCCTCACATCCCAGAAGAAGGTGGAGCTTGTCCTTAAGCAGCT 2539
Qy 2295 GCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCT 2354
Db 2540 TCGATTAATGCAATCATCTCAAAGTATGTCCAAGCTCACTCTGACTCCATGGGTTGGTCT 2599
Qy 2355 TCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTT 2414
Db 2600 TCGGAAGATCAATGTGTCTTACTGGTGCTGGGAGGATATGTCTCCATTCACAAATAGTTT 2659
Qy 2415 ACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACC 2474
Db 2660 GCTGCAGTGGATGCCATCTGAGCCCAGTGATGCTGGCTTCTGTGGGATCTTGTCAGAGCC 2719
Qy 2475 CAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGA 2534
Db 2720 TAGTACTCGGGGATTAAAGGCTGCAACCTGCATCAACCCTCTCAATGGCAGCGTCTGTGA 2779
Qy 2535 AAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATG 2594
Db 2780 AAGGCCTGCAAACCACAGTGCCAAGCAGTGCCGGACACCATGTGCCCTGCGGACAGCGTG 2839
Qy 2595 TGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGT 2654
Db 2840 TGGCGAGTGCACTAGCAGCAGCTCGGAGTGCATGTGGTGCAGTAACATGAAGCAGTGTGT 2899
Qy 2655 GGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGAT 2714
Db 2900 GGACTCCAATGCCTACGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGAT 2959
Qy 2715 GAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCA 2774
Db 2960 GAGCAGCTGCCCACCTGAAAATTGCTCTGGCTACTGTACCTGCAGCCATTGCTTGGAGCA 3019
Qy 2775 ACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGG 2834
Db 3020 GCCAGGCTGTGGTTGGTGTACTGATCCTAGCAATACTGGGAAAGGAAAATGTATTGAGGG 3079
Qy 2835 TTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACA 2894
Db 3080 CAGCTATAAAGGACCTGTGAAGATGCCGTCACAGGCCTCTGCAGGAAATGTGTATCCACA 3139
Qy 2895 GCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCA 2954
Db 3140 GCCCCTTCTGAACTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCA 3199
Qy 2955 CTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAA 3014
Db 3200 CTGTCCAGCTTGCCAGTGCAACGGACACAGCAAATGCATCAACCAGAGTGTCTGTGAGAA 3259
Qy 3015 GTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGA 3074
Db 3260 GTGTGAGGACCTGACCACGGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTATGGTGA 3319
Qy 3075 TCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAA 3134
Db 3320 CCCGACTAATGGAGGCAAATGTCAGCCATGCAAGTGCAATGGGCACGCATCACTGTGCAA 3379
Qy 3135 CACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCT 3194
Db 3380 CACCAACACCGGCAAGTGCTTCTGTACCACCAAAGGTGTCAAGGGGGACGAGTGCCAGCT 3439
Qy 3195 ATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCT 3254
Db 3440 ATGTAAGGTAGAAAATCGATACCAGGGGAACCCTCTCAAAGGAACATGCTACTATACCCT 3499
Qy 3255 TCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGC 3314
Db 3500 TCTCATTGACTATCAGTTCACCTTTAGCCTGTCCCAGGGAGACGACCGCTACTACACAGC 3559
Qy 3315 TATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGC 3374
Db 3560 CATCAACTTTGTGGCTACTCCTGATGAACAAAACAGGGATTTTGACATGTTCATCAATGC 3619

```
Qy       3375 CTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGC 3434
Db       3620 CTCCAAAAAATTCAACCTCAACATCACCTGGGCCACCAGCTTCCCAGCCGGAACCCAGAC 3679

Qy       3435 TGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTC 3494
Db       3680 TGGAGAAGAGGTGCCTGTTGTTTCAAAAACCAACATCAAGGAATACAAAGATAGCTTCTC 3739

Qy       3495 TAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAA 3554
Db       3740 TAATGAGAAATTTGATTTTCGCAACCATCCAAACATCACTTTCTTTGTTTATGTCAGTAA 3799

Qy       3555 TTTCACCTGGCCCATCAAAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCT 3614
Db       3800 TTTCACTTGGCCCATCAAAATTCAGATTGCCTTCTCCCAGCACAGCAACTTTATGGACCT 3859

Qy       3615 GGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGT 3674
Db       3860 GGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCGCTGCTTCTGGTGGCTGCAGT 3919

Qy       3675 GGTTTGGAAGATCAAACAAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGA 3734
Db       3920 GGTCTGGAAGATCAAGCAGAGCTGTTGGGCATCCAGGCGGAGAGAGCAACTTCTTCGGGA 3979

Qy       3735 GATGCAACAGATGGCCAGCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGA 3794
Db       3980 GATGCAACAGATGGCCAGCCGCCCCTTTGCTTCTGTAAACGTTGCCTTGGAAACAGATGA 4039

Qy       3795 GGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGA 3854
Db       4040 AGAACCTCCTGATCTCATTGGGGGAAGTATAAAGACCGTTCCTAAGCCCATTGCCCTGGA 4099

Qy       3855 GCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCT 3914
Db       4100 GCCCTGCTTTGGTAACAAAGCCGCAGTCCTCTCTGTATTCGTGAGGCTCCCTCGAGGACT 4159

Qy       3915 GGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACAT 3974
Db       4160 GGGAGGAATCCCTCCTCCTGGTCAGTCAGGTCTCGCTGTGGCCAGTGCCCTGGTGGACAT 4219

Qy       3975 TTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCA 4034
Db       4220 TTCTCAGCAGATGCCAATAGTGTACAAGGAGAAGTCAGGAGCTGTAAGAAACCGGAAGCA 4279

Qy       4035 GCAGCCCCCTGCACAGCCTGGGACCTGCATCTGA 4068
Db       4280 GCAGCCGCCTGCACAGCCTGGAACCTGCATTTGA 4313
```

```
RESULT 9
AB038387
LOCUS       AB038387               8739 bp    mRNA    linear   ROD 18-JAN-2001
DEFINITION  Rattus norvegicus atrn mRNA for attractin, complete cds.
ACCESSION   AB038387
VERSION     AB038387.1  GI:12275307
KEYWORDS    attractin.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (sites)
  AUTHORS   Kuramoto,T., Kitada,K., Inui,T., Sasaki,Y., Ito,K., Hase,T.,
            Kawaguchi,S., Ogawa,Y., Nakao,K., Barsh,G.S., Nagao,M., Ushijima,T.
            and Serikawa,T.
  TITLE     Attractin/mahogany/zitter plays a critical role in myelination of
            the central nervous system
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (2), 559-564 (2001)
  MEDLINE   21143347
   PUBMED   11209055
REFERENCE   2  (bases 1 to 8739)
  AUTHORS   Kuramoto,T., Serikawa,T. and Ushijima,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-FEB-2000) Takashi Kuramoto, National Cancer Center
            Research Institute, Carcinogenesis Division; Tsukiji 5-1-1,
            Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramot@gan2.ncc.go.jp,
            Tel:81-33542-2511(ex.4521), Fax:81-35565-1753)
FEATURES             Location/Qualifiers
     source          1. .8739
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="WTC"
                     /db_xref="taxon:10116"
     gene            1. .8739
                     /gene="atrn"
     CDS             61. .4359
                     /gene="atrn"
                     /codon_start=1
                     /product="attractin"
                     /protein_id="BAB21017.1"
                     /db_xref="GI:12275308"
                     /translation="MVAAAAAAEATEARLRGYTTATAAPAGWKERQHRPCAATGAWRP
                     WPRAGLCLPRVLSRALSPPPLLPLLPLLFSLLLLPLPREAEAAAVAAAVSGSAAAEAK
                     ECDRPCVNGGRCNPGTGQCVCPTGWVGEQCQHCGGRFRLTGSSGFVTDGPGNYKYKTK
                     CTWLIEGQPNKIMRLRFNHFATECSWDHLYVYDGDSIYAPLIAAFSGLIVPERDGNET
                     APEVTVTSGYALLHFFSDAAYNLTGFNITYNFDMCPNNCSGRGECKSSNSSSTVECEC
                     SENWKGESCDIPHCTDNCGFPHRGICNASDTRGCSCFPHWQGPGCSIPVPANQSFWTR
                     EEYSDLKLPRASHKAVVNGNIMWVVGGYMFNHSDYSMVLAYDLASREWLSLNHSVNSV
                     VVRYGHSLALHKDKIYMYGGKIDSTGNVTNELRVFHIHNESWVLLTPKAKDQYAVVGH
                     SAHIVTLSSGRVVMLVIFGHCPLYGYISVVQEYDLEKNTWSILQTQGALVQGGYGHSS
                     VYDHRTKALYVHGGYKAFSANKYRLADDLYRYHVDTQMWTILKDSRFFRYLHTAVIVS
                     GTMLVFGGNTHNDTSMSHGAKCFSSDFMAYDIACDRWSVLPRPELHHDVNRFGHSAVL
                     HNSTMYVFGGFNSLLLSDVLVFTSEQCDAHRSEAACVAAGPGIRCLWDTQSSRCTSWE
                     LATEEQAEKLKSECFSKRTLDHDRCDQHTDCYSCTANTNDCHWCNDHCVPVNHSCTEG
                     QISIAKYDNCPKDNPMYYCNKKTSCRSCALDQNCQWEPRNQECIALPENICGIGWHLV
                     GNSCLKITTAKENYDNAKLSCRNHNAFLASLTSQKKVEFVLKQLRLMQSSQSTSKLTL
                     TPWVGLRKINVSYWCWEDMSPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCIN
                     PLNGSVCERPANHSAKQCRTPCALRTACGECTSSSSECMWCSNMKQCVDSNAYVASFP
                     FGQCMEWYTMSSCPPENCSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKM
                     PSHASTGNVYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCEDLTT
                     GKHCETCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGEECQLCEVE
                     NRYQGNPLKGTCYYTLLIDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASK
                     NFNLNITWATSFPAGTQTGEEVPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSN
                     FTWPIKIQIAFSQHSNFMDLVQFFVTFFSCFLSLLLVAAVVWKIKQSCWASRRREQLL
                     REMQQMASRPFASVNVALETDEEPPDLIGGSIKTVPKPIALEPCFGNKAAVLSVFVRL
                     PRGLGGIPPPGQSGLAVASALVDISQQMPIVYKEKSGAVRNRKQQPPAQPGTCI"
ORIGIN

  Query Match           82.4%;  Score 3353.2;  DB 10;  Length 8739;
  Best Local Similarity 89.2%;  Pred. No. 0;
  Matches 3616;  Conservative    0;  Mismatches  438;  Indels    0;  Gaps    0;

Qy         15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74
Db        306 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCTGTGTCAGGCTCGGCCGCAGCTGAGGCCAA 365

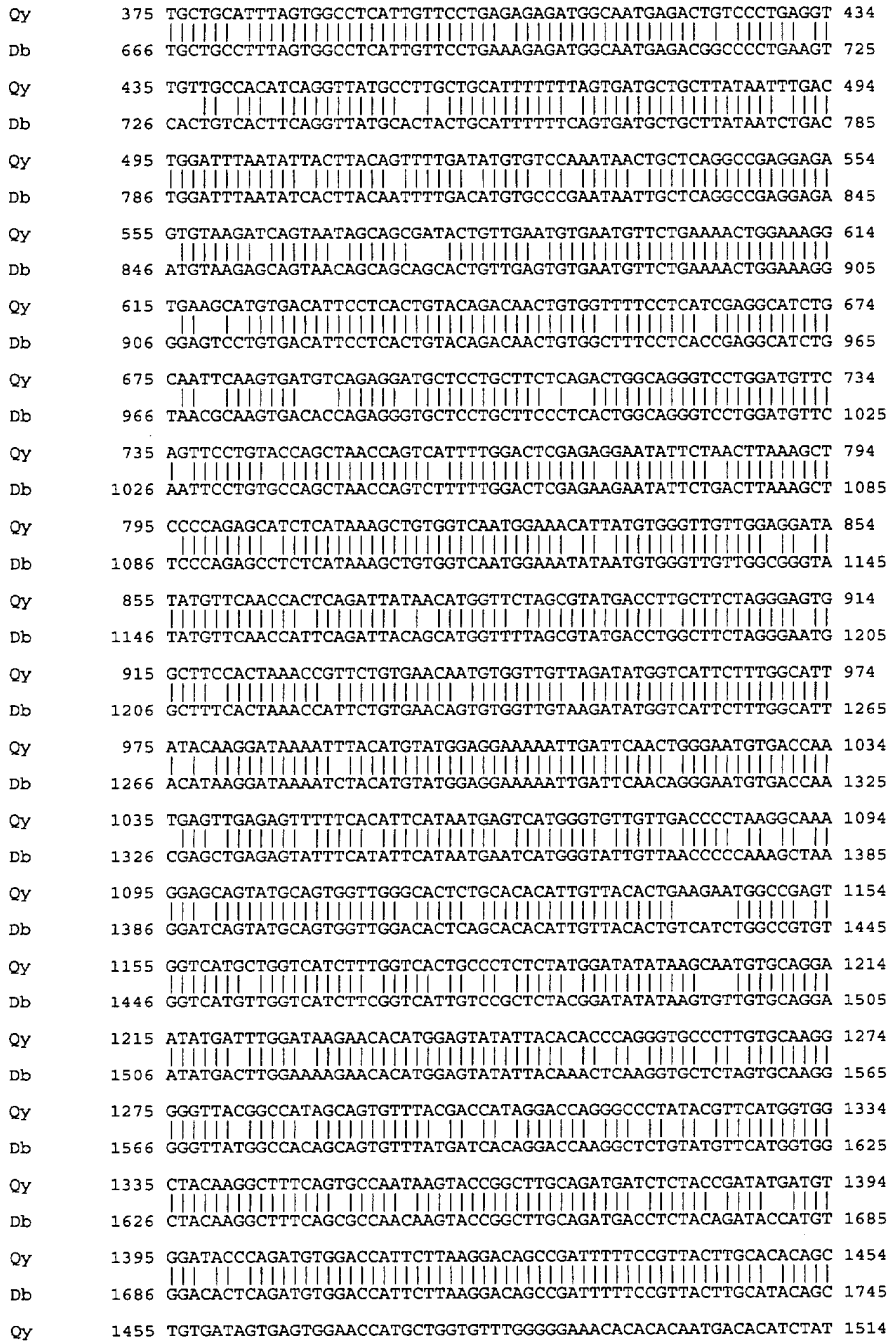
Qy         75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134
Db        366 GGAATGTGACCGGCCGTGTGTCAACGGCGGCCGCTGCAACCCTGGCACCGGCCAGTGCGT 425

Qy        135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194
Db        426 CTGCCCCACGGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 485

Qy        195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254
Db        486 TGGATCTTCTGGATTTGTAACAGATGGACCTGGGAATTATAAATATAAAACGAAATGCAC 545

Qy        255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314
Db        546 ATGGCTCATTGAAGGACAGCCAAATAAGATAATGAGACTTCGATTCAACCATTTTGCTAC 605

Qy        315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374
Db        606 TGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATTTACGCACCTCTGAT 665
```


Qy 375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434
Db 666 TGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCCCCTGAAGT 725
Qy 435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494
Db 726 CACTGTCACTTCAGGTTATGCACTACTGCATTTTTTCAGTGATGCTGCTTATAATCTGAC 785
Qy 495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554
Db 786 TGGATTTAATATCACTTACAATTTTGACATGTGCCCGAATAATTGCTCAGGCCGAGGAGA 845
Qy 555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614
Db 846 ATGTAAGAGCAGTAACAGCAGCAGCACTGTTGAGTGTGAATGTTCTGAAAACTGGAAAGG 905
Qy 615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 674
Db 906 GGAGTCCTGTGACATTCCTCACTGTACAGACAACTGTGGCTTTCCTCACCGAGGCATCTG 965
Qy 675 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 734
Db 966 TAACGCAAGTGACACCAGAGGGTGCTCCTGCTTCCCTCACTGGCAGGGTCCTGGATGTTC 1025
Qy 735 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 794
Db 1026 AATTCCTGTGCCAGCTAACCAGTCTTTTTGGACTCGAGAAGAATATTCTGACTTAAAGCT 1085
Qy 795 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 854
Db 1086 TCCCAGAGCCTCTCATAAAGCTGTGGTCAATGGAAATATAATGTGGGTTGTTGGCGGGTA 1145
Qy 855 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 914
Db 1146 TATGTTCAACCATTCAGATTACAGCATGGTTTTAGCGTATGACCTGGCTTCTAGGGAATG 1205
Qy 915 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 974
Db 1206 GCTTTCACTAAACCATTCTGTGAACAGTGTGGTTGTAAGATATGGTCATTCTTTGGCATT 1265
Qy 975 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1034
Db 1266 ACATAAGGATAAAATCTACATGTATGGAGGAAAAATTGATTCAACAGGGAATGTGACCAA 1325
Qy 1035 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1094
Db 1326 CGAGCTGAGAGTATTTCATATTCATAATGAATCATGGGTATTGTTAACCCCCAAAGCTAA 1385
Qy 1095 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1154
Db 1386 GGATCAGTATGCAGTGGTTGGACACTCAGCACACATTGTTACACTGTCATCTGGCCGTGT 1445
Qy 1155 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1214
Db 1446 GGTCATGTTGGTCATCTTCGGTCATTGTCCGCTCTACGGATATATAAGTGTTGTGCAGGA 1505
Qy 1215 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1274
Db 1506 ATATGACTTGGAAAAGAACACATGGAGTATATTACAAACTCAAGGTGCTCTAGTGCAAGG 1565
Qy 1275 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1334
Db 1566 GGGTTATGGCCACAGCAGTGTTTATGATCACAGGACCAAGGCTCTGTATGTTCATGGTGG 1625
Qy 1335 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1394
Db 1626 CTACAAGGCTTTCAGCGCCAACAAGTACCGGCTTGCAGATGACCTCTACAGATACCATGT 1685
Qy 1395 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1454
Db 1686 GGACACTCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCATACAGC 1745
Qy 1455 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1514
Db 1746 TGTGATAGTGAGTGGAACCATGCTGGTGTTCGGAGGGAACACACACAATGACACTTCCAT 1805
Qy 1515 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1574
Db 1806 GAGCCACGGTGCCAAATGCTTCTCCTCAGACTTCATGGCTTATGACATTGCTTGTGACCG 1865
Qy 1575 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGC 1634
Db 1866 ATGGTCAGTTCTCCCCAGACCTGAGCTCCATCACGATGTCAACCGATTTGGCCACTCAGC 1925
Qy 1635 AGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGA 1694
Db 1926 AGTCTTGCACAACAGCACCATGTATGTGTTCGGCGGCTTCAACAGCCTCCTCCTCAGTGA 1985
Qy 1695 CATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGC 1754
Db 1986 CGTCTTAGTCTTCACCTCGGAGCAGTGTGATGCACACCGCAGCGAAGCTGCTTGTGTGGC 2045
Qy 1755 AGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGC 1814
Db 2046 AGCAGGACCTGGCATCCGGTGTCTGTGGGACACACAGTCGTCTCGATGTACCTCCTGGGA 2105
Qy 1815 GCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCT 1874
Db 2106 GTTGGCAACTGAAGAACAAGCAGAAAAGTTAAAATCAGAATGTTTCTCTAAAAGAACCCT 2165
Qy 1875 TGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGA 1934
Db 2166 TGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAATACCAATGA 2225
Qy 1935 CTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGAT 1994
Db 2226 CTGCCACTGGTGCAATGATCACTGTGTCCCTGTGAACCACAGCTGCACAGAAGGCCAGAT 2285
Qy 1995 CTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAA 2054
Db 2286 CTCCATTGCCAAGTATGACAATTGCCCCAAGGATAACCCCATGTACTACTGCAATAAGAA 2345
Qy 2055 GACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGA 2114
Db 2346 AACCAGCTGCAGGAGCTGTGCCCTAGATCAGAATTGCCAGTGGGAACCTCGAAATCAAGA 2405
Qy 2115 GTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATG 2174
Db 2406 GTGCATCGCCCTGCCCGAAAACATCTGTGGCATTGGCTGGCATTTGGTCGGAAACTCGTG 2465
Qy 2175 TTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCA 2234
Db 2466 TCTGAAAATCACTACTGCTAAGGAAAATTATGACAATGCTAAATTGTCCTGTAGAAACCA 2525
Qy 2235 CAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCT 2294
Db 2526 CAATGCCTTTTTGGCTTCGCTTACGTCCCAGAAGAAGGTGGAGTTTGTCCTTAAGCAGCT 2585
Qy 2295 GCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCT 2354
Db 2586 TCGATTAATGCAATCATCTCAGAGCACGTCCAAGCTCACTCTGACTCCATGGGTTGGTCT 2645
Qy 2355 TCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTT 2414
Db 2646 TCGGAAGATCAATGTATCTTACTGGTGCTGGGAGGATATGTCTCCATTCACAAATAGTTT 2705
Qy 2415 ACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACC 2474
Db 2706 GCTGCAGTGGATGCCATCTGAGCCCAGTGATGCCGGCTTCTGTGGGATCTTGTCAGAGCC 2765
Qy 2475 CAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGA 2534
Db 2766 TAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGAAGTGTTTGTGA 2825
Qy 2535 AAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATG 2594

```
Db       2826 AAGGCCTGCAAACCACAGTGCCAAGCAGTGCCGGACACCCTGTGCCCTGCGGACAGCATG 2885

Qy       2595 TGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGT 2654

Db       2886 TGGTGAGTGCACCAGCAGCAGCTCCGAATGCATGTGGTGCAGCAACATGAAGCAGTGCGT 2945

Qy       2655 GGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGAT 2714

Db       2946 AGACTCCAATGCCTACGTAGCCTCCTTTCCTTTCGGCCAGTGTATGGAGTGGTATACAAT 3005

Qy       2715 GAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCA 2774

Db       3006 GAGCAGCTGCCCACCTGAAAATTGCTCTGGCTACTGTACCTGCAGCCATTGCTTGGAGCA 3065

Qy       2775 ACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGG 2834

Db       3066 GCCAGGCTGTGGTTGGTGTACTGATCCTAGCAATACTGGGAAAGGCAAATGTATTGAGGG 3125

Qy       2835 TTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACA 2894

Db       3126 AAGCTATAAAGGACCCGTGAAGATGCCTTCACATGCCTCTACAGGAAATGTGTACCCACA 3185

Qy       2895 GCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCA 2954

Db       3186 GCCCCTTCTGAACTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTTATTCA 3245

Qy       2955 CTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAA 3014

Db       3246 CTGTCCAGCTTGCCAGTGCAATGGACACAGCAAGTGCATCAACCAGAGCATCTGTGAGAA 3305

Qy       3015 GTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGA 3074

Db       3306 GTGTGAGGACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTATGGTGA 3365

Qy       3075 TCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAA 3134

Db       3366 CCCGACTAATGGAGGCAAATGTCAGCCATGCAAGTGCAACGGGCACGCATCACTGTGCAA 3425

Qy       3135 CACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCT 3194

Db       3426 CACCAACACTGGCAAGTGCTTCTGTACCACCAAGGGCGTCAAGGGAGAGGAGTGCCAGCT 3485

Qy       3195 ATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCT 3254

Db       3486 ATGTGAGGTAGAAAATCGATACCAGGGAAACCCTCTCAAAGGAACATGTTACTATACCCT 3545

Qy       3255 TCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGC 3314

Db       3546 TCTCATTGACTATCAGTTCACTTTTAGCCTGTCCCAGGAAGACGACCGCTATTACACAGC 3605

Qy       3315 TATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGC 3374

Db       3606 CATCAATTTTGTGGCTACTCCTGATGAACAAAACAGGGATTTGGACATGTTCATCAATGC 3665

Qy       3375 CTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGC 3434

Db       3666 CTCCAAAAACTTCAACCTCAACATCACCTGGGCCACCAGTTTCCCAGCTGGAACCCAGAC 3725

Qy       3435 TGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTC 3494

Db       3726 TGGAGAAGAGGTGCCTGTTGTTTCAAAAACCAACATCAAGGAATACAAAGATAGCTTCTC 3785

Qy       3495 TAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAA 3554

Db       3786 TAATGAGAAATTTGATTTTCGCAACCATCCAAACATTACTTTCTTTGTTTATGTCAGTAA 3845

Qy       3555 TTTCACCTGGCCCATCAAAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCT 3614

Db       3846 TTTCACTTGGCCCATCAAAATTCAAATTGCCTTCTCCCAGCACAGCAATTTTATGGACCT 3905

Qy       3615 GGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGT 3674

Db       3906 GGTGCAGTTCTTCGTGACTTTCTTCAGCTGTTTCCTCTCTCTGCTTCTGGTGGCGGCAGT 3965

Qy       3675 GGTTTGGAAGATCAAACAAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGA 3734

Db       3966 GGTCTGGAAGATCAAGCAGAGCTGTTGGGCATCCAGGAGGAGAGAGCAACTTCTTCGGGA 4025

Qy       3735 GATGCAACAGATGGCCAGCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGA 3794

Db       4026 GATGCAACAGATGGCCAGCCGCCCTTTTGCTTCTGTAAATGTTGCCTTGGAAACAGATGA 4085

Qy       3795 GGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGA 3854

Db       4086 AGAACCTCCTGATCTCATTGGGGGAAGTATAAAGACTGTTCCTAAGCCCATTGCCCTGGA 4145

Qy       3855 GCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCT 3914

Db       4146 GCCCTGCTTTGGTAACAAAGCTGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGACT 4205

Qy       3915 GGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACAT 3974

Db       4206 GGGAGGAATCCCTCCTCCTGGTCAGTCAGGTCTTGCTGTGGCCAGTGCCCTGGTGGACAT 4265

Qy       3975 TTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCA 4034

Db       4266 TTCTCAGCAGATGCCAATAGTGTACAAGGAGAAGTCAGGAGCTGTAAGGAACCGGAAGCA 4325

Qy       4035 GCAGCCCCCTGCACAGCCTGGGACCTGCATCTGA 4068

Db       4326 GCAGCCCCCTGCACAGCCTGGAACCTGCATTTGA 4359
```

RESULT 10
AR164807
LOCUS AR164807 8827 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 1 from patent US 6274339.
ACCESSION AR164807
VERSION AR164807.1 GI:16238010
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 8827)
AUTHORS Moore,K. and Nagle,D.Lynn.
TITLE Methods and compositions for the diagnosis and treatment of body weight disorders, including obesity
JOURNAL Patent: US 6274339-A 1 14-AUG-2001;
FEATURES Location/Qualifiers
source 1. .8827
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN

Query Match 81.8%; Score 3326.4; DB 6; Length 8827;
Best Local Similarity 91.1%; Pred. No. 0;
Matches 3534; Conservative 0; Mismatches 346; Indels 0; Gaps 0;

```
Qy        189 ACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAA 248

Db        677 ACTAACTGGCTCTTCTGGATTTGTAACAGATGGACCTGGGAATTATAAATATAAGACGAA 736

Qy        249 GTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTT 308

Db        737 GTGCACATGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGCTTCAACCATTT 796

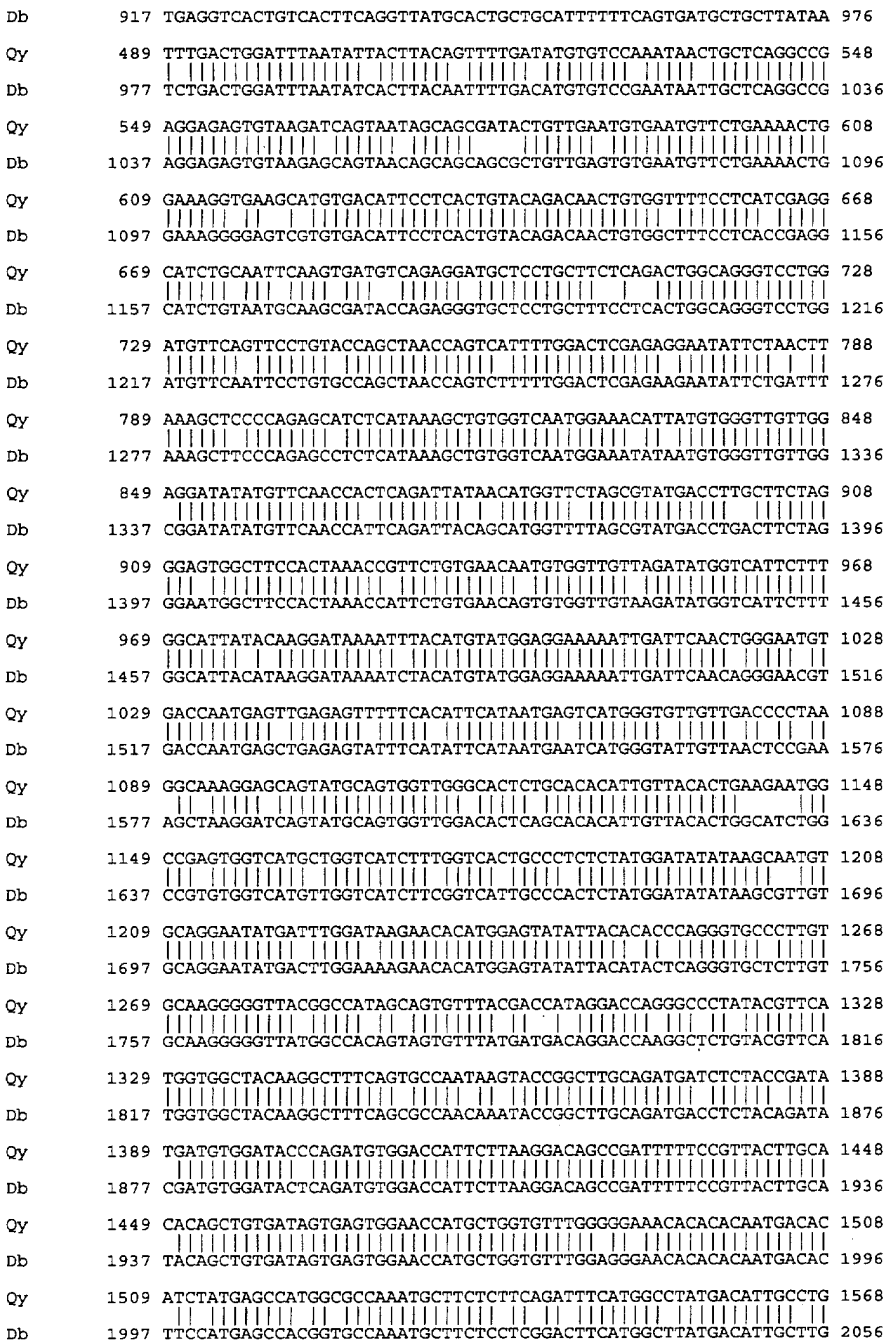
Qy        309 TGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACC 368

Db        797 TGCTACAGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATCTACGCACC 856

Qy        369 GCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCC 428

Db        857 TCTGATTGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCTCC 916

Qy        429 TGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAA 488
```


Db 917 TGAGGTCACTGTCACTTCAGGTTATGCACTGCTGCATTTTTTCAGTGATGCTGCTTATAA 976
Qy 489 TTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCG 548
Db 977 TCTGACTGGATTTAATATCACTTACAATTTTGACATGTGTCCGAATAATTGCTCAGGCCG 1036
Qy 549 AGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTG 608
Db 1037 AGGAGAGTGTAAGAGCAGTAACAGCAGCAGCGCTGTTGAGTGTGAATGTTCTGAAAACTG 1096
Qy 609 GAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGG 668
Db 1097 GAAAGGGGAGTCGTGTGACATTCCTCACTGTACAGACAACTGTGGCTTTCCTCACCGAGG 1156
Qy 669 CATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGG 728
Db 1157 CATCTGTAATGCAAGCGATACCAGAGGGTGCTCCTGCTTTCCTCACTGGCAGGGTCCTGG 1216
Qy 729 ATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTT 788
Db 1217 ATGTTCAATTCCTGTGCCAGCTAACCAGTCTTTTTGGACTCGAGAAGAATATTCTGATTT 1276
Qy 789 AAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGG 848
Db 1277 AAAGCTTCCCAGAGCCTCTCATAAAGCTGTGGTCAATGGAAATATAATGTGGGTTGTTGG 1336
Qy 849 AGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAG 908
Db 1337 CGGATATATGTTCAACCATTCAGATTACAGCATGGTTTTAGCGTATGACCTGACTTCTAG 1396
Qy 909 GGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTT 968
Db 1397 GGAATGGCTTCCACTAAACCATTCTGTGAACAGTGTGGTTGTAAGATATGGTCATTCTTT 1456
Qy 969 GGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGT 1028
Db 1457 GGCATTACATAAGGATAAAATCTACATGTATGGAGGAAAAATTGATTCAACAGGGAACGT 1516
Qy 1029 GACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAA 1088
Db 1517 GACCAATGAGCTGAGAGTATTTCATATTCATAATGAATCATGGGTATTGTTAACTCCGAA 1576
Qy 1089 GGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGG 1148
Db 1577 AGCTAAGGATCAGTATGCAGTGGTTGGACACTCAGCACACATTGTTACACTGGCATCTGG 1636
Qy 1149 CCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGT 1208
Db 1637 CCGTGTGGTCATGTTGGTCATCTTCGGTCATTGCCCACTCTATGGATATATAAGCGTTGT 1696
Qy 1209 GCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGT 1268
Db 1697 GCAGGAATATGACTTGGAAAAGAACACATGGAGTATATTACATACTCAGGGTGCTCTTGT 1756
Qy 1269 GCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCA 1328
Db 1757 GCAAGGGGGTTATGGCCACAGTAGTGTTTATGATGACAGGACCAAGGCTCTGTACGTTCA 1816
Qy 1329 TGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATA 1388
Db 1817 TGGTGGCTACAAGGCTTTCAGCGCCAACAAATACCGGCTTGCAGATGACCTCTACAGATA 1876
Qy 1389 TGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCA 1448
Db 1877 CGATGTGGATACTCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCA 1936
Qy 1449 CACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACAC 1508
Db 1937 TACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGAGGGAACACACACAATGACAC 1996
Qy 1509 ATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTG 1568
Db 1997 TTCCATGAGCCACGGTGCCAAATGCTTCTCCTCGGACTTCATGGCTTATGACATTGCTTG 2056
Qy 1569 TGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCA 1628
Db 2057 TGACCGATGGTCAGTGCTTCCCAGACCTGAGCTCCATCATGATGTCAACAGATTTGGCCA 2116
Qy 1629 TTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCT 1688
Db 2117 TTCAGCAGTCTTGTACAACAGCACCATGTATGTGTTCGGCGGCTTCAACAGCCTCCTCCT 2176
Qy 1689 CAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTG 1748
Db 2177 CAGTGACGTCTTGGTCTTTACCTCGGAGCAGTGCGATGCACACCGCAGTGAAGCTGCTTG 2236
Qy 1749 TTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTC 1808
Db 2237 TGTGGCAGCAGGACCTGGTATCCGGTGTCTGTGGGACACACAGTCGTCTCGATGTACCTC 2296
Qy 1809 GTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAG 1868
Db 2297 CTGGGAGTTGGCAACTGAAGAACAAGCAGAAAAGTTAAAATCAGAGTGTTTTTCTAAAAG 2356
Qy 1869 AACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACAC 1928
Db 2357 AACCCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAATAC 2416
Qy 1929 CAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGG 1988
Db 2417 CAATGACTGCCACTGGTGCAATGATCACTGTGTCCCTGTGAACCACAGCTGCACAGAAGG 2476
Qy 1989 CCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAA 2048
Db 2477 CCAGATCTCCATTGCCAAGTATGAGAGTTGCCCCAAGGATAACCCCATGTACTACTGCAA 2536
Qy 2049 CAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAA 2108
Db 2537 TAAGAAAACCAGCTGCAGGAGCTGTGCCCTAGACCAGAACTGCCAGTGGGAGCCCCGGAA 2596
Qy 2109 TCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAA 2168
Db 2597 TCAAGAGTGCATCGCCCTGCCGGAAAATATCTGTGGCAATGGCTGGCATTTGGTTGGAAA 2656
Qy 2169 CTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAG 2228
Db 2657 CTCGTGTCTGAAAATCACTACTGCTAAGGAGAATTATGACAATGCTAAATTGTCCTGTAG 2716
Qy 2229 GAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAA 2288
Db 2717 GAACCACAATGCCTTTTTGGCTTCCCTCACATCCCAGAAGAAGGTGGAGTTTGTCCTTAA 2776
Qy 2289 GCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGT 2348
Db 2777 GCAGCTTCGATTAATGCAATCATCTCAAAGTATGTCCAAGCTCACTCTGACTCCATGGGT 2836
Qy 2349 CGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAA 2408
Db 2837 TGGTCTTCGGAAGATCAATGTGTCTTACTGGTGCTGGGAGGATATGTCTCCATTCACAAA 2896
Qy 2409 TAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAAATTTTATC 2468
Db 2897 TAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGTGATGCTGGCTTCTGTGGGATCTTGTC 2956
Qy 2469 AGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGT 2528
Db 2957 AGAGCCTAGTACTCGGGGATTAAAGGCTGCAACCTGCATCAACCCTCTCAATGGCAGCGT 3016
Qy 2529 CTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGAC 2588
Db 3017 CTGTGAAAGGCCTGCAAACCACAGTGCCAAGCAGTGCCGGACACCATGTGCCCTGCGGAC 3076
Qy 2589 AGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCA 2648
Db 3077 AGCGTGTGGCGAGTGCACTAGCAGCAGCTCGGAGTGCATGTGGTGCAGTAACATGAAGCA 3136

```
Qy       2649 GTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTA 2708
             |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db       3137 GTGTGTGGACTCCAATGCCTACGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTA 3196

Qy       2709 TACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTT 2768
             |||||||||||| |||||| |||||||||||| || |||||||||||||| || ||||||||
Db       3197 TACGATGAGCAGCTGCCCACCTGAAAATTGCTCTGGCTACTGTACCTGCAGCCATTGCTT 3256

Qy       2769 GGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCAT 2828
             |||||| |||||||||| |||||||||||||| |||||||||| ||||| ||||| ||
Db       3257 GGAGCAGCCAGGCTGTGGTTGGTGTACTGATCCTAGCAATACTGGGAAAGGAAAATGTAT 3316

Qy       2829 AGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTA 2888
              |||||  |||||||||||| |||||||||| || || ||| || |||||||| | ||
Db       3317 TGAGGGCAGCTATAAAGGACCTGTGAAGATGCCGTCACAGGCCTCTGCAGGAAATGTGTA 3376

Qy       2889 TCCACAGCCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTT 2948
             |||||||||||| || || ||||||||||||||||||||||||||||||||||||||||||
Db       3377 TCCACAGCCCCTTCTGAACTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTT 3436

Qy       2949 CATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTG 3008
             |||||||||||||||||||| |||||||| ||||| |||||||||| ||||| |||||
Db       3437 CATTCACTGTCCAGCTTGCCAGTGCAACGGACACAGCAAATGCATCAACCAGAGTATCTG 3496

Qy       3009 TGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTA 3068
             |||||||||||| ||||||||| ||||||||||||||||||||||||||||||||||||||
Db       3497 TGAGAAGTGTGAGGACCTGACCACGGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTA 3556

Qy       3069 CGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCT 3128
             ||||| || || |||||||||||||||||||||||||||||||||||||||||||| || ||
Db       3557 TGGTGACCCGACTAATGGAGGCAAATGTCAGCCATGCAAGTGCAATGGGCACGCATCACT 3616

Qy       3129 GTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTG 3188
             |||||||||||||| |||||||||||||| |||||||| || ||||||||||||||||||
Db       3617 GTGCAACACCAACACCGGCAAGTGCTTCTGTACCACCAAAGGTGTCAAGGGGGACGAGTG 3676

Qy       3189 CCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTA 3248
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||| || ||
Db       3677 CCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAAAGGAACATGCTACTA 3736

Qy       3249 TACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTA 3308
             ||| ||||| |||||||||||||||||||||| || |||||||||| || ||||| ||
Db       3737 TACCCTTCTCATTGACTATCAGTTCACCTTTAGCCTGTCCCAGGAAGACGACCGCTACTA 3796

Qy       3309 CACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCAT 3368
             |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3797 CACAGCCATCAACTTTGTGGCTACTCCTGATGAACAAAACAGGGATTTGGACATGTTCAT 3856

Qy       3369 CAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAAC 3428
             |||||||||||| || |||||||||||||||||||||  |||| ||| |||| |||||
Db       3857 CAATGCCTCCAAAAACTTCAACCTCAACATCACCTGGGCCACCAGCTTCCCAGCCGGAAC 3916

Qy       3429 CCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAG 3488
             |||| |||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db       3917 CCAGACTGGAGAAGAGGTGCCTGTTGTTTCAAAAACCAACATCAAGGAATACAAAGATAG 3976

Qy       3489 TTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGT 3548
             |||||||||||||| |||||||||||||||||||| ||||| ||||||||||||||||||
Db       3977 CTTCTCTAATGAGAAATTTGATTTTCGCAACCATCCAAACATCACTTTCTTTGTTTATGT 4036

Qy       3549 CAGTAATTTCACCTGGCCCATCAAAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTAT 3608
             |||||||||||| |||||||||||||||||||||||||||||||| |||||||||| || ||
Db       4037 CAGTAATTTCACTTGGCCCATCAAAATTCAGATTGCCTTCTCCCAGCACAGCAACTTCAT 4096

Qy       3609 GGACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGC 3668
             |||||||||||||||||||||||||||||||||||||||||||||||  |||| ||||||||
Db       4097 GGACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCGCTGCTTCTGGTGGC 4156

Qy       3669 TGCTGTGGTTTGGAAGATCAAACAAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCT 3728
             ||| ||||| |||||||||||| || || |||||||| ||||| || ||||||||||||||
Db       4157 TGCAGTGGTCTGGAAGATCAAGCAGAGCTGTTGGGCATCCAGGCGGAGAGAGCAACTTCT 4216

Qy       3729 TCGAGAGATGCAACAGATGGCCAGCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAAC 3788
             ||| |||||||||||||||||||||| |||||||| |||||||| || ||||||||||||
Db       4217 TCGGGAGATGCAACAGATGGCCAGCCGCCCCTTTGCTTCTGTAAACGTTGCCTTGGAAAC 4276

Qy       3789 AGATGAGGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGC 3848
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4277 AGATGAGGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGC 4336

Qy       3849 ACTGGAGCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCG 3908
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4337 ACTGGAGCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCG 4396

Qy       3909 AGGCCTGGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGT 3968
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4397 AGGCCTGGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGT 4456

Qy       3969 GGACATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCG 4028
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4457 GGACATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCG 4516

Qy       4029 GAAGCAGCAGCCCCCTGCACAGCCTGGGACCTGCATCTGA 4068
             |||||||||||||||||||||||||||||||||||||||||
Db       4517 GAAGCAGCAGCCCCCTGCACAGCCTGGGACCTGCATCTGA 4556
```

RESULT 11
AB038388

```
LOCUS       AB038388                4500 bp    mRNA    linear   ROD 18-JAN-2001
DEFINITION  Rattus norvegicus Atrn mRNA for attractin, complete cds.
ACCESSION   AB038388
VERSION     AB038388.1  GI:12275311
KEYWORDS    attractin.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (sites)
  AUTHORS   Kuramoto,T., Kitada,K., Inui,T., Sasaki,Y., Ito,K., Hase,T.,
            Kawaguchi,S., Ogawa,Y., Nakao,K., Barsh,G.S., Nagao,M., Ushijima,T.
            and Serikawa,T.
  TITLE     Attractin/mahogany/zitter plays a critical role in myelination of
            the central nervous system
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (2), 559-564 (2001)
  MEDLINE   21143347
   PUBMED   11209055
REFERENCE   2  (bases 1 to 4500)
  AUTHORS   Kuramoto,T., Serikawa,T. and Ushijima,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-FEB-2000) Takashi Kuramoto, National Cancer Center
            Research Institute, Carcinogenesis Division; Tsukiji 5-1-1,
            Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramot@gan2.ncc.go.jp,
            Tel:81-33542-2511(ex.4521), Fax:81-35565-1753)
FEATURES             Location/Qualifiers
     source          1. .4500
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="WTC"
                     /db_xref="taxon:10116"
     gene            1. .4500
                     /gene="Atrn"
     CDS             61. .3888
                     /gene="Atrn"
                     /codon_start=1
                     /product="attractin"
                     /protein_id="BAB21018.1"
                     /db_xref="GI:12275312"
                     /translation="MVAAAAAAEATEARLRGYTTATAAPAGWKERQHRPCAATGAWRP
                     WPRAGLCLPRVLSRALSPPPLLPLLPLLFSLLLLPLPREAEAAAVAAAVSGSAAAEAK
                     ECDRPCVNGGRCNPGTGQCVCPTGWVGEQCQHCGGRFRLTGSSGFVTDGPGNYKYKTK
                     CTWLIEGQPNKIMRLRFNHFATECSWDHLYVYDGDSIYAPLIAAFSGLIVPERDGNET
                     APEVTVTSGYALLHFFSDAAYNLTGFNITYNFDMCPNNCSGRGECKSSNSSSTVECEC
                     SENWKGESCDIPHCTDNCGFPHRGICNASDTRGCSCFPHWQGPGCSIPVPANQSFWTR
                     EEYSDLKLPRASHKAEVNGNIMWVVGGYMFNHSDYSMVLAYDLASREWLSLNHSVNSV
```

VVRYGHSLALHKDKIYMYGGKIDSTGNVTNELRVFHIHNESWVLLTPKAKDQYAVVGH
SAHIVTLSSGRVVMLVIFGHCPLYGYISVVQEYDLEKNTWSILQTQGALVQGGYGHSS
VYDHRTKALYVHGGYKAFSANKYRLADDLYRYHVDTQMWTILKDSRFFRYLHTAVIVS
GTMLVFGGNTHNDTSMSHGAKCFSSDFMAYDIACDRWSVLPRPELHHDVNRFGHSAVL
HNSTMYVFGGFNSLLLSDVLVFTSEQCDAHRSEAACVAAGPGIRCLWDTQSSRCTSWE
LATEEQAEKLKSECFSKRTLDHDRCDQHTDCYSCTANTNDCHWCNDHCVPVNHSCTEG
QISIAKYDNCPKDNPMYYCNKKTSCRSCALDQNCQWEPRNQECIALPENICGIGWHLV
GNSCLKITTAKENYDNAKLSCRNHNAFLASLTSQKKVEFVLKQLRLMQSSQSTSKLTL
TPWVGLRKINVSYWCWEDMSPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCIN
PLNGSVCERPANHSAKQCRTPCALRTACGECTSSSSECMWCSNMKQCVDSNAYVASFP
FGQCMEWYTMSSCPPENCSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKM
PSHASTGNVYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCEDLTT
GKHCETCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGEECQLCEVE
NRYQGNPLKGTCYYTLLIDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASK
NFNLNITWATSFPAGTQTGEEVPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSN
FTWPIKIQVRVTS"

ORIGIN

Query Match 72.0%; Score 2928.8; DB 10; Length 4500;
Best Local Similarity 88.9%; Pred. No. 0;
Matches 3167; Conservative 0; Mismatches 397; Indels 0; Gaps 0;

```
Qy      15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74
Db     306 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCTGTGTCAGGCTCGGCCGCAGCTGAGGCCAA 365

Qy      75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134
Db     366 GGAATGTGACCGGCCGTGTGTCAACGGCGGCCGCTGCAACCCTGGCACCGGGCAGTGCGT 425

Qy     135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194
Db     426 CTGCCCCACGGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 485

Qy     195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254
Db     486 TGGATCTTCTGGATTTGTAACAGATGGACCTGGGAATTATAAATATAAAACGAAATGCAC 545

Qy     255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314
Db     546 ATGGCTCATTGAAGGACAGCCAAATAAGATAATGAGACTTCGATTCAACCATTTTGCTAC 605

Qy     315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374
Db     606 TGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATTTACGCACCTCTGAT 665

Qy     375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434
Db     666 TGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCCCCTGAAGT 725

Qy     435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494
Db     726 CACTGTCACTTCAGGTTATGCACTACTGCATTTTTTCAGTGATGCTGCTTATAATCTGAC 785

Qy     495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554
Db     786 TGGATTTAATATCACTTACAATTTTGACATGTGCCCGAATAATTGCTCAGGCCGAGGAGA 845

Qy     555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614
Db     846 ATGTAAGAGCAGTAACAGCAGCAGCACTGTTGAGTGTGAATGTTCTGAAAACTGGAAAGG 905

Qy     615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 674
Db     906 GGAGTCCTGTGACATTCCTCACTGTACAGACAACTGTGGCTTTCCTCACCGAGGCATCTG 965

Qy     675 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 734
Db     966 TAACGCAAGTGACACCAGAGGGTGCTCCTGCTTCCCTCACTGGCAGGGTCCTGGATGTTC 1025

Qy     735 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 794
Db    1026 AATTCCTGTGCCAGCTAACCAGTCTTTTTGGACTCGAGAAGAATATTCTGACTTAAAGCT 1085

Qy     795 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 854
Db    1086 TCCCAGAGCCTCTCATAAAGCTGAGGTCAATGGAAATATAATGTGGGTTGTTGGCGGGTA 1145

Qy     855 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 914
Db    1146 TATGTTCAACCATTCAGATTACAGCATGGTTTTAGCGTATGACCTGGCTTCTAGGGAATG 1205

Qy     915 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 974
Db    1206 GCTTTCACTAAACCATTCTGTGAACAGTGTGGTTGTAAGATATGGTCATTCTTTGGCATT 1265

Qy     975 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1034
Db    1266 ACATAAGGATAAAATCTACATGTATGGAGGAAAAATTGATTCAACAGGGAATGTGACCAA 1325

Qy    1035 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1094
Db    1326 CGAGCTGAGAGTATTTCATATTCATAATGAATCATGGGTATTGTTAACCCCCAAAGCTAA 1385

Qy    1095 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1154
Db    1386 GGATCAGTATGCAGTGGTTGGACACTCAGCACACATTGTTACACTGTCATCTGGCCGTGT 1445

Qy    1155 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1214
Db    1446 GGTCATGTTGGTCATCTTCGGTCATTGTCCGCTCTACGGATATATAAGTGTTGTGCAGGA 1505

Qy    1215 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1274
Db    1506 ATATGACTTGGAAAAGAACACATGGAGTATATTACAAACTCAAGGTGCTCTAGTGCAAGG 1565

Qy    1275 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1334
Db    1566 GGGTTATGGCCACAGCAGTGTTTATGATCACAGGACCAAGGCTCTGTATGTTCATGGTGG 1625

Qy    1335 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1394
Db    1626 CTACAAGGCTTTCAGCGCCAACAAGTACCGGCTTGCAGATGACCTCTACAGATACCATGT 1685

Qy    1395 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1454
Db    1686 GGACACTCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCATACAGC 1745

Qy    1455 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1514
Db    1746 TGTGATAGTGAGTGGAACCATGCTGGTGTTCGGAGGGAACACACACAATGACACTTCCAT 1805

Qy    1515 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1574
Db    1806 GAGCCACGGTGCCAAATGCTTCTCCTCAGACTTCATGGCTTATGACATTGCTTGTGACCG 1865

Qy    1575 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGC 1634
Db    1866 ATGGTCAGTTCTCCCCAGACCTGAGCTCCATCACGATGTCAACCGATTTGGCCACTCAGC 1925

Qy    1635 AGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGA 1694
Db    1926 AGTCTTGCACAACAGCACCATGTATGTGTTCGGCGGCTTCAACAGCCTCCTCCTCAGTGA 1985

Qy    1695 CATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGC 1754
Db    1986 CGTCTTAGTCTTCACCTCGGAGCAGTGTGATGCACACCGCAGCGAAGCTGCTTGTGTGGC 2045

Qy    1755 AGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGC 1814
Db    2046 AGCAGGACCTGGCATCCGGTGTCTGTGGGACACACAGTCGTCTCGATGTACCTCCTGGGA 2105

Qy    1815 GCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCT 1874
Db    2106 GTTGGCAACTGAAGAACAAGCAGAAAAGTTAAAATCAGAATGTTTCTCTAAAAGAACCCT 2165
```

```
Qy       1875 TGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGA 1934
Db       2166 TGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAATACCAATGA 2225

Qy       1935 CTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGAT 1994
Db       2226 CTGCCACTGGTGCAATGATCACTGTGTCCCTGTGAACCACAGCTGCACAGAAGGCCAGAT 2285

Qy       1995 CTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAA 2054
Db       2286 CTCCATTGCCAAGTATGACAATTGCCCCAAGGATAACCCCATGTACTACTGCAATAAGAA 2345

Qy       2055 GACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGA 2114
Db       2346 AACCAGCTGCAGGAGCTGTGCCCTAGATCAGAATTGCCAGTGGGAACCTCGAAATCAAGA 2405

Qy       2115 GTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATG 2174
Db       2406 GTGCATCGCCCTGCCCGAAAACATCTGTGGCATTGGCTGGCATTTGGTCGGAAACTCGTG 2465

Qy       2175 TTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCA 2234
Db       2466 TCTGAAAATCACTACTGCTAAGGAAAATTATGACAATGCTAAATTGTCCTGTAGAAACCA 2525

Qy       2235 CAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCT 2294
Db       2526 CAATGCCTTTTTGGCTTCGCTTACGTCCCAGAAGAAGGTGGAGTTTGTCCTTAAGCAGCT 2585

Qy       2295 GCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCT 2354
Db       2586 TCGATTAATGCAATCATCTCAGAGCACGTCCAAGCTCACTCTGACTCCATGGGTTGGTCT 2645

Qy       2355 TCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTT 2414
Db       2646 TCGGAAGATCAATGTATCTTACTGGTGCTGGGAGGATATGTCTCCATTCACAAATAGTTT 2705

Qy       2415 ACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACC 2474
Db       2706 GCTGCAGTGGATGCCATCTGAGCCCAGTGATGCCGGCTTCTGTGGGATCTTGTCAGAGCC 2765

Qy       2475 CAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGA 2534
Db       2766 TAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGAAGTGTTTGTGA 2825

Qy       2535 AAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATG 2594
Db       2826 AAGGCCTGCAAACCACAGTGCCAAGCAGTGCCGGACACCCTGTGCCCTGCGGACAGCATG 2885

Qy       2595 TGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGT 2654
Db       2886 TGGTGAGTGCACCAGCAGCAGCTCCGAATGCATGTGGTGCAGCAACATGAAGCAGTGCGT 2945

Qy       2655 GGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGAT 2714
Db       2946 AGACTCCAATGCCTACGTAGCCTCCTTTCCTTTCGGCCAGTGTATGGAGTGGTATACAAT 3005

Qy       2715 GAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCA 2774
Db       3006 GAGCAGCTGCCCACCTGAAAATTGCTCTGGCTACTGTACCTGCAGCCATTGCTTGGAGCA 3065

Qy       2775 ACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGG 2834
Db       3066 GCCAGGCTGTGGTTGGTGTACTGATCCTAGCAATACTGGGAAAGGCAAATGTATTGAGGG 3125

Qy       2835 TTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACA 2894
Db       3126 AAGCTATAAAGGACCCGTGAAGATGCCTTCACATGCCTCTACAGGAAATGTGTACCCACA 3185

Qy       2895 GCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCA 2954
Db       3186 GCCCCTTCTGAACTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTTATTCA 3245

Qy       2955 CTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAA 3014
Db       3246 CTGTCCAGCTTGCCAGTGCAATGGACACAGCAAGTGCATCAACCAGAGCATCTGTGAGAA 3305

Qy       3015 GTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGA 3074
Db       3306 GTGTGAGGACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTATGGTGA 3365

Qy       3075 TCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAA 3134
Db       3366 CCCGACTAATGGAGGCAAATGTCAGCCATGCAAGTGCAACGGGCACGCATCACTGTGCAA 3425

Qy       3135 CACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCT 3194
Db       3426 CACCAACACTGGCAAGTGCTTCTGTACCACCAAGGGCGTCAAGGGAGAGGAGTGCCAGCT 3485

Qy       3195 ATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCT 3254
Db       3486 ATGTGAGGTAGAAAATCGATACCAGGGAAACCCTCTCAAAGGAACATGTTACTATACCCT 3545

Qy       3255 TCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGC 3314
Db       3546 TCTCATTGACTATCAGTTCACTTTTAGCCTGTCCCAGGAAGACGACCGCTATTACACAGC 3605

Qy       3315 TATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGC 3374
Db       3606 CATCAATTTTGTGGCTACTCCTGATGAACAAAACAGGGATTTGGACATGTTCATCAATGC 3665

Qy       3375 CTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGC 3434
Db       3666 CTCCAAAAACTTCAACCTCAACATCACCTGGGCCACCAGTTTCCCAGCTGGAACCCAGAC 3725

Qy       3435 TGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTC 3494
Db       3726 TGGAGAAGAGGTGCCTGTTGTTTCAAAAACCAACATCAAGGAATACAAAGATAGCTTCTC 3785

Qy       3495 TAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAA 3554
Db       3786 TAATGAGAAATTTGATTTTCGCAACCATCCAAACATTACTTTCTTTGTTTATGTCAGTAA 3845

Qy       3555 TTTCACCTGGCCCATCAAAATTCA 3578
Db       3846 TTTCACTTGGCCCATCAAAATTCA 3869
```

RESULT 12
AB093245

LOCUS	AB093245 5683 bp mRNA linear ROD 07-FEB-2003
DEFINITION	Mus musculus mRNA for mKIAA0548 protein.
ACCESSION	AB093245
VERSION	AB093245.1 GI:26006172
KEYWORDS	FLI_CDNA.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1
AUTHORS	Okazaki,N., Kikuno,R., Ohara,R., Inamoto,S., Hara,Y., Nagase,T., Ohara,O. and Koga,H.
TITLE	Prediction of the coding sequences of mouse homologues of KIAA gene: I. The complete nucleotide sequences of 100 mouse KIAA-homologous cDNAs identified by screening of terminal sequences of cDNA clones randomly sampled from size-fractionated libraries
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 5683)
AUTHORS	Okazaki,N., Kikuno,R., Nagase,T., Ohara,O. and Koga,H.
TITLE	Direct Submission
JOURNAL	Submitted (04-OCT-2002) Hisashi Koga, Kazusa DNA Research Institute, Laboratory for Genome Informatics; 2-6-7 Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan (E-mail:mouse@kazusa.or.jp, Tel:81-438-52-3919, Fax:81-438-52-3918)
COMMENT	The CREATE program supported by Japan science and technology corporation; cDNA full insert sequencing: Kazusa DNA Research Institute; cDNA library construction, clone selection and 5'- &

```
          3'-end one pass sequencing.
FEATURES             Location/Qualifiers
     source          1. .5683
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="mbg03041"
                     /tissue_type="brain"
                     /dev_stage="adult"
                     /note="vector:modified pBC SK+"
     gene            1. .5683
                     /gene="mKIAA0548"
     CDS             <1. .3181
                     /gene="mKIAA0548"
                     /note="CDS is predicted by in silico analysis. Start codon
                     is not identified."
                     /codon_start=2
                     /evidence=not_experimental
                     /product="mKIAA0548 protein"
                     /protein_id="BAC41429.1"
                     /db_xref="GI:26006173"
                     /translation="AYDLTSREWLPLNHSVNSVVVRYGHSLALHKDKIYMYGGKIDST
                     GNVTNELRVFHIHNESWVLLTPKAKDQYAVVGHSAHIVTLASGRVVMLVIFGHCPLYG
                     YISVVQEYDLEKNTWSILHTQGALVQGGYGHSSVYDDRTKALYVHGGYKAFSANKYRL
                     ADDLYRYDVDTQMWTILKDSRFFRYLHTAVIVSGTMLVFGGNTHNDTSMSHGAKCFSS
                     DFMAYDIACDRWSVLPRPELHHDVNRFGHSAVLYNSTMYVFGGFNSLLLSDVLVFTLE
                     QCDAHRSEAACVAAGPGIRCLWDTQSSRCTSWELATEEQAEKLKSECFSKRTLDHDRC
                     DQHTDCYSCTANTNDCHWCNDHCVPVNHSCTEGQISIAKYESCPKDNPMYYCNKKTSC
                     RSCALDQNCQWEPRNQECIALPENICGNGWHLVGNSCLKITTAKENYDNAKLSCRNHN
                     AFLASLTSQKKVEFVLKQLRLMQSSQSMSKLTLTPWVGLRKINVSYWCWEDMSPFTNS
                     LLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLNGSVCERPANHSAKQCRTPCALR
                     TACGECTSSSSECMWCSNMKQCVDSNAYVASFPFGQCMEWYTMSSCPPENCSGYCTCS
                     HCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKMPSQASAGNVYPQPLLNSSMCLEDSR
                     YNWSFIHCPACQCNGHSKCINQSVCEKCEDLTTGKHCETCISGFYGDPTNGGKCQPCK
                     CNGHASLCNTNTGKCFCTTKGVKGDECQLCEVENRYQGNPLKGTCYYTLLIDYQFTFS
                     LSQEDDRYYTAINFVATPDEQNRDLDMFINASKNFNLNITWATSFPAGTQTGEEVPVV
                     SKTTIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTWPIKIQIAFSQHSNFMDLVQFFV
                     TFFSCFLSLLLVAAVVWKIKQSCWASRRREQLLREMQQMASRPFASVNVALETDEEPP
                     DLIGGSIKTVPKPIALEPCFGNKAAVLSVFVRLPRGLGGIPPPGQSGLAVASALVDIS
                     QQMPIVYKEKSGAVRNRKQQPPAQPGTCI"
ORIGIN

  Query Match             66.2%;  Score 2694.6;  DB 10;  Length 5683;
  Best Local Similarity   90.4%;  Pred. No. 0;
  Matches 2877;  Conservative    0;  Mismatches  304;  Indels    0;  Gaps    0;

Qy        888 AGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGT 947
              ||||||||||||  ||||||||| |||||||||||||||| |||||||||| ||||||
Db          1 AGCGTATGACCTGACTTCTAGGGAATGGCTTCCACTAAACCATTCTGTGAACAGTGTGGT 60

Qy        948 TGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAA 1007
              ||| |||||||||||||||||||||| | |||||||||| |||||||||||||||||||
Db         61 TGTAAGATATGGTCATTCTTTGGCATTACATAAGGATAAAATCTACATGTATGGAGGAAA 120

Qy       1008 AATTGATTCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTC 1067
              |||||||||||| ||||| |||||||||||| ||||||| ||||| |||||||||| ||
Db        121 AATTGATTCAACAGGGAACGTGACCAATGAGCTGAGAGTATTTCATATTCATAATGAATC 180

Qy       1068 ATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACA 1127
              |||||| ||||| || || || || |||||| |||||||||||||||| ||||| |||||
Db        181 ATGGGTATTGTTAACTCCGAAAGCTAAGGATCAGTATGCAGTGGTTGGACACTCAGCACA 240

Qy       1128 CATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCT 1187
              |||||||||||||      |||||| ||||||||| |||||||||| ||||| ||||| ||
Db        241 CATTGTTACACTGGCATCTGGCCGTGTGGTCATGTTGGTCATCTTCGGTCATTGCCCACT 300

Qy       1188 CTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATT 1247
              |||||||||||||||  |||||||||||||| ||||| ||||||||||||||||||||||
Db        301 CTATGGATATATAAGCGTTGTGCAGGAATATGACTTGGAAAAGAACACATGGAGTATATT 360

Qy       1248 ACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAG 1307
              ||| || |||||||| |||||||||||||||| ||||| || |||||||| ||  | ||
Db        361 ACATACTCAGGGTGCTCTTGTGCAAGGGGGTTATGGCCACAGTAGTGTTTATGATGACAG 420

Qy       1308 GACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCT 1367
              ||||| ||| || |||||||||||||||||||||||||||||| ||||| || ||||||||
Db        421 GACCAAGGCTCTGTACGTTCATGGTGGCTACAAGGCTTTCAGCGCCAACAAATACCGGCT 480

Qy       1368 TGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAG 1427
              |||||||| |||||| |||| |||||||||| |||||||||||||||||||||||||||
Db        481 TGCAGATGACCTCTACAGATACGATGTGGATACTCAGATGTGGACCATTCTTAAGGACAG 540

Qy       1428 CCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGG 1487
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        541 CCGATTTTTCCGTTACTTGCATACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGG 600

Qy       1488 GGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTT 1547
              || |||||||||||||||| || |||||||| || |||||||||||||| || || ||
Db        601 AGGGAACACACACAATGACACTTCCATGAGCCACGGTGCCAAATGCTTCTCCTCGGACTT 660

Qy       1548 CATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCA 1607
              |||||| |||||||||| |||||||| |||||||||||||||||||| ||||| ||
Db        661 CATGGCTTATGACATTGCTTGTGACCGATGGTCAGTGCTTCCCAGACCTGAGCTCCATCA 720

Qy       1608 TGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGG 1667
              |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db        721 TGATGTCAACAGATTTGGCCATTCAGCAGTCTTGTACAACAGCACCATGTATGTGTTCGG 780

Qy       1668 TGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGC 1727
              || ||||| || |||||||||| ||| || |||| || |||| ||| ||||| |||||
Db        781 CGGCTTCAACAGCCTCCTCCTCAGTGACGTCTTGGTCTTTACCTTGGAGCAGTGCGATGC 840

Qy       1728 GCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACAC 1787
              || || |||||||| |||||| | |||||||||||||||||| |||||| |||| ||||
Db        841 ACACCGCAGTGAAGCTGCTTGTGTGGCAGCAGGACCTGGTATCCGGTGTCTGTGGGACAC 900

Qy       1788 AGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAA 1847
              | |||||||| |||| ||| |||| | |||||||||| ||||||| ||||||||||||
Db        901 ACAGTCGTCTCGATGTACCTCCTGGGAGTTGGCAACTGAAGAACAAGCAGAAAAGTTAAA 960

Qy       1848 ATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTG 1907
              |||||| |||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        961 ATCAGAGTGTTTTTCTAAAAGAACCCTTGACCATGACAGATGTGACCAGCACACAGATTG 1020

Qy       1908 TTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAG 1967
              |||||||||||||||| |||||||||||||||||||||||||||| || ||||||||
Db        1021 TTACAGCTGCACAGCCAATACCAATGACTGCCACTGGTGCAATGATCACTGTGTCCCTGT 1080

Qy       1968 GAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGA 2027
              |||||||||||| |||||||||||||||||||    | |||||||| ||||||||||||
Db        1081 GAACCACAGCTGCACAGAAGGCCAGATCTCCATTGCCAAGTATGAGAGTTGCCCCAAGGA 1140

Qy       2028 TAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAA 2087
              |||||| |||||||||| || ||||| |||||||||||||||||||||||||| |||||
Db        1141 TAACCCCATGTACTACTGCAATAAGAAAACCAGCTGCAGGAGCTGTGCCCTAGACCAGAA 1200

Qy       2088 CTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCAT 2147
              |||||||||||||||||||||||| |||||||| |||||||| ||||||||||||||||
Db        1201 CTGCCAGTGGGAGCCCCGGAATCAAGAGTGCATCGCCCTGCCGGAAAATATCTGTGGCAA 1260

Qy       2148 TGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGA 2207
              |||||||||||||||||||||||| |||||||| |||||||||||||||||||||||||
Db        1261 TGGCTGGCATTTGGTTGGAAACTCGTGTCTGAAAATCACTACTGCTAAGGAGAATTATGA 1320

Qy       2208 CAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAA 2267
              |||||||||||| |||||||||||||||||||| |||||||| || ||| |||||||
Db        1321 CAATGCTAAATTGTCCTGTAGGAACCACAATGCCTTTTTGGCTTCCCTCACATCCCAGAA 1380

Qy       2268 GAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAA 2327
              |||||| || |||||||||||||||| ||| ||||||| |||||||| || ||||||||
Db        1381 GAAGGTGGAGTTTGTCCTTAAGCAGCTTCGATTAATGCAATCATCTCAAAGTATGTCCAA 1440

Qy       2328 GCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGA 2387
              |||||| | || |||||||| || |||||||||||||||||||||||| ||||||||||
Db        1441 GCTCACTCTGACTCCATGGGTTGGTCTTCGGAAGATCAATGTGTCTTACTGGTGCTGGGA 1500
```

```
Qy       2388 AGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGC 2447
Db       1501 GGATATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGTGATGC 1560

Qy       2448 TGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCAT 2507
Db       1561 TGGCTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGATTAAAGGCTGCAACCTGCAT 1620

Qy       2508 CAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCG 2567
Db       1621 CAACCCTCTCAATGGCAGCGTCTGTGAAAGGCCTGCAAACCACAGTGCCAAGCAGTGCCG 1680

Qy       2568 GACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCAT 2627
Db       1681 GACACCATGTGCCCTGCGGACAGCGTGTGGCGAGTGCACTAGCAGCAGCTCGGAGTGCAT 1740

Qy       2628 GTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTT 2687
Db       1741 GTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTACGTGGCCTCCTTCCCTTT 1800

Qy       2688 TGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTA 2747
Db       1801 TGGCCAGTGTATGGAATGGTATACGATGAGCAGCTGCCCACCTGAAAATTGCTCTGGCTA 1860

Qy       2748 CTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAA 2807
Db       1861 TTGTACCTGCAGCCATTGCTTGGAGCAGCCAGGCTGTGGTTGGTGTACTGATCCTAGCAA 1920

Qy       2808 TACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCA 2867
Db       1921 TACTGGGAAAGGAAAATGTATTGAGGGCAGCTATAAAGGACCTGTGAAGATGCCGTCACA 1980

Qy       2868 AGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGA 2927
Db       1981 GGCCTCTGCAGGAAATGTGTATCCACAGCCCCTTCTGAACTCCAGCATGTGTCTAGAGGA 2040

Qy       2928 CAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAA 2987
Db       2041 CAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAGTGCAACGGACACAGCAA 2100

Qy       2988 ATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGA 3047
Db       2101 ATGCATCAACCAGAGTGTCTGTGAGAAGTGTGAGGACCTGACCACGGGCAAGCACTGCGA 2160

Qy       3048 GACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAA 3107
Db       2161 GACCTGCATATCTGGCTTCTATGGTGACCCGACTAATGGAGGCAAATGTCAGCCATGCAA 2220

Qy       3108 GTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAA 3167
Db       2221 GTGCAATGGGCACGCATCACTGTGCAACACCAACACCGGCAAGTGCTTCTGTACCACCAA 2280

Qy       3168 GGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCC 3227
Db       2281 AGGTGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAGGGAAATCC 2340

Qy       3228 TCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATC 3287
Db       2341 TCTCAAAGGAACATGCTACTATACCCTTCTCATTGACTATCAGTTCACCTTTAGCCTGTC 2400

Qy       3288 CCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAA 3347
Db       2401 CCAGGAAGACGACCGCTACTACACAGCCATCAACTTTGTGGCTACTCCTGATGAACAAAA 2460

Qy       3348 CAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGC 3407
Db       2461 CAGGGATTTGGACATGTTCATCAATGCCTCCAAAAACTTCAACCTCAACATCACCTGGGC 2520

Qy       3408 TGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAA 3467
Db       2521 CACCAGCTTCCCAGCCGGAACCCAGACTGGAGAAGAGGTGCCTGTTGTTTCAAAAACCAC 2580

Qy       3468 CATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAA 3527
Db       2581 CATCAAGGAATACAAAGATAGCTTCTCTAATGAGAAATTTGATTTTCGCAACCATCCAAA 2640

Qy       3528 TATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGATTGCCTT 3587
Db       2641 CATCACTTTCTTTGTTTATGTCAGTAATTTCACTTGGCCCATCAAAATTCAGATTGCCTT 2700

Qy       3588 CTCTCAGCACAGCAATTTTATGGACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTT 3647
Db       2701 CTCCCAGCACAGCAACTTCATGGACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTT 2760

Qy       3648 CCTCTCTTTGCTCCTGGTGGCTGCTGTGGTTTGGAAGATCAAACAAAGTTGTTGGGCCTC 3707
Db       2761 CCTCTCGCTGCTTCTGGTGGCTGCAGTGGTCTGGAAGATCAAGCAGAGCTGTTGGGCATC 2820

Qy       3708 CAGACGTAGAGAGCAACTTCTTCGAGAGATGCAACAGATGGCCAGCCGTCCCTTTGCCTC 3767
Db       2821 CAGGCGGAGAGAGCAACTTCTTCGGGAGATGCAACAGATGGCCAGCCGCCCCTTTGCTTC 2880

Qy       3768 TGTAAATGTCGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTATTGGGGGGAGTATAAA 3827
Db       2881 TGTAAACGTTGCCTTGGAAACAGATGAAGAACCTCCTGATCTCATTGGGGGAAGTATAAA 2940

Qy       3828 GACTGTTCCCAAACCCATTGCACTGGAGCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTC 3887
Db       2941 GACCGTTCCTAAGCCCATTGCCCTGGAGCCCTGCTTTGGTAACAAAGCCGCAGTCCTCTC 3000

Qy       3888 TGTGTTTGTGAGGCTCCCTCGAGGCCTGGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCT 3947
Db       3001 TGTATTCGTGAGGCTCCCTCGAGGACTGGGAGGAATCCCTCCTCCTGGTCAGTCAGGTCT 3060

Qy       3948 TGCTGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAA 4007
Db       3061 CGCTGTGGCCAGTGCCCTGGTGGACATTTCTCAGCAGATGCCAATAGTGTACAAGGAGAA 3120

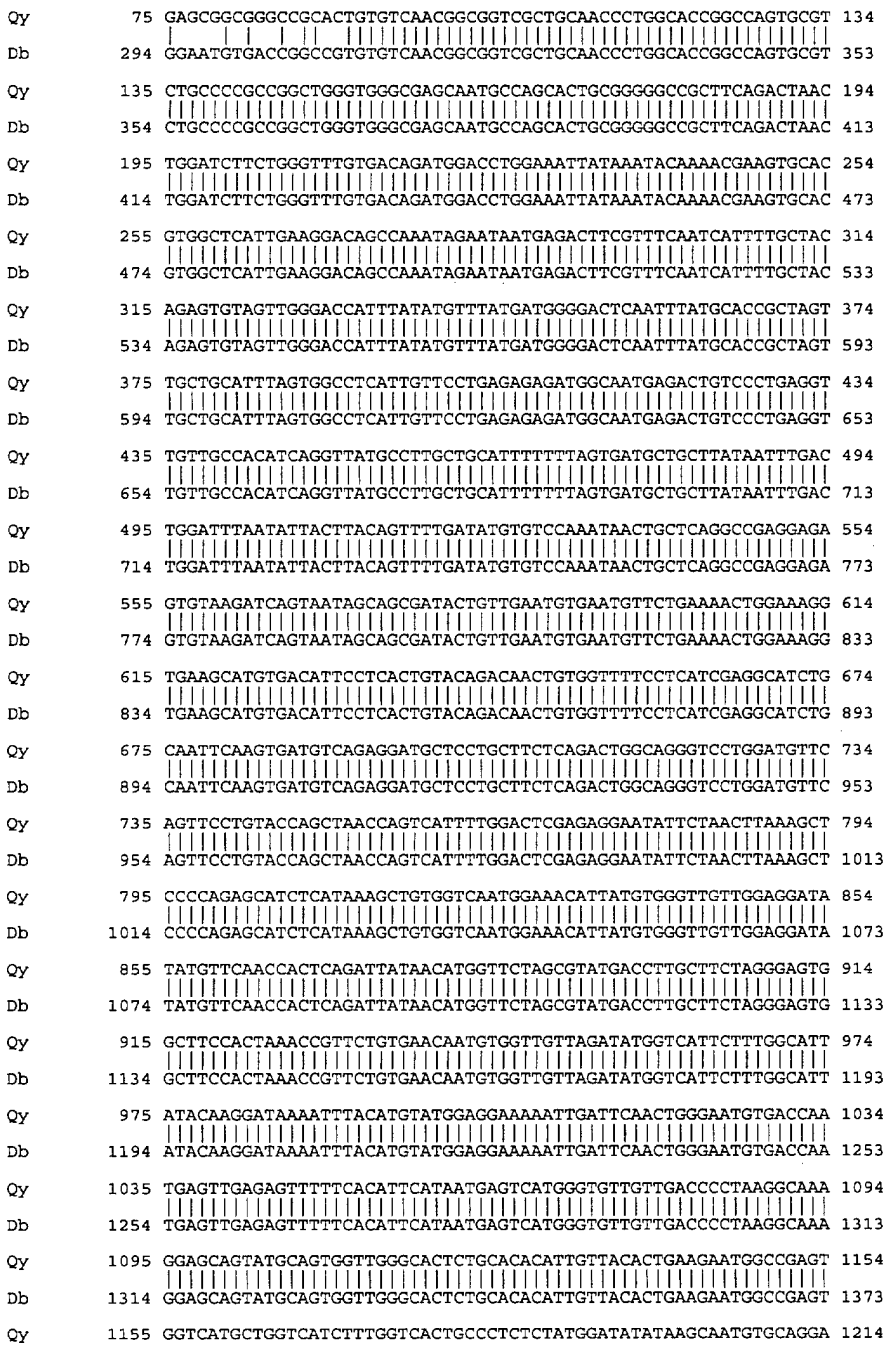
Qy       4008 GTCAGGAGCCGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGCCTGGGACCTGCATCTG 4067
Db       3121 GTCAGGAGCTGTAAGAAACCGGAAGCAGCAGCCGCCTGCACAGCCTGGAACCTGCATTTG 3180

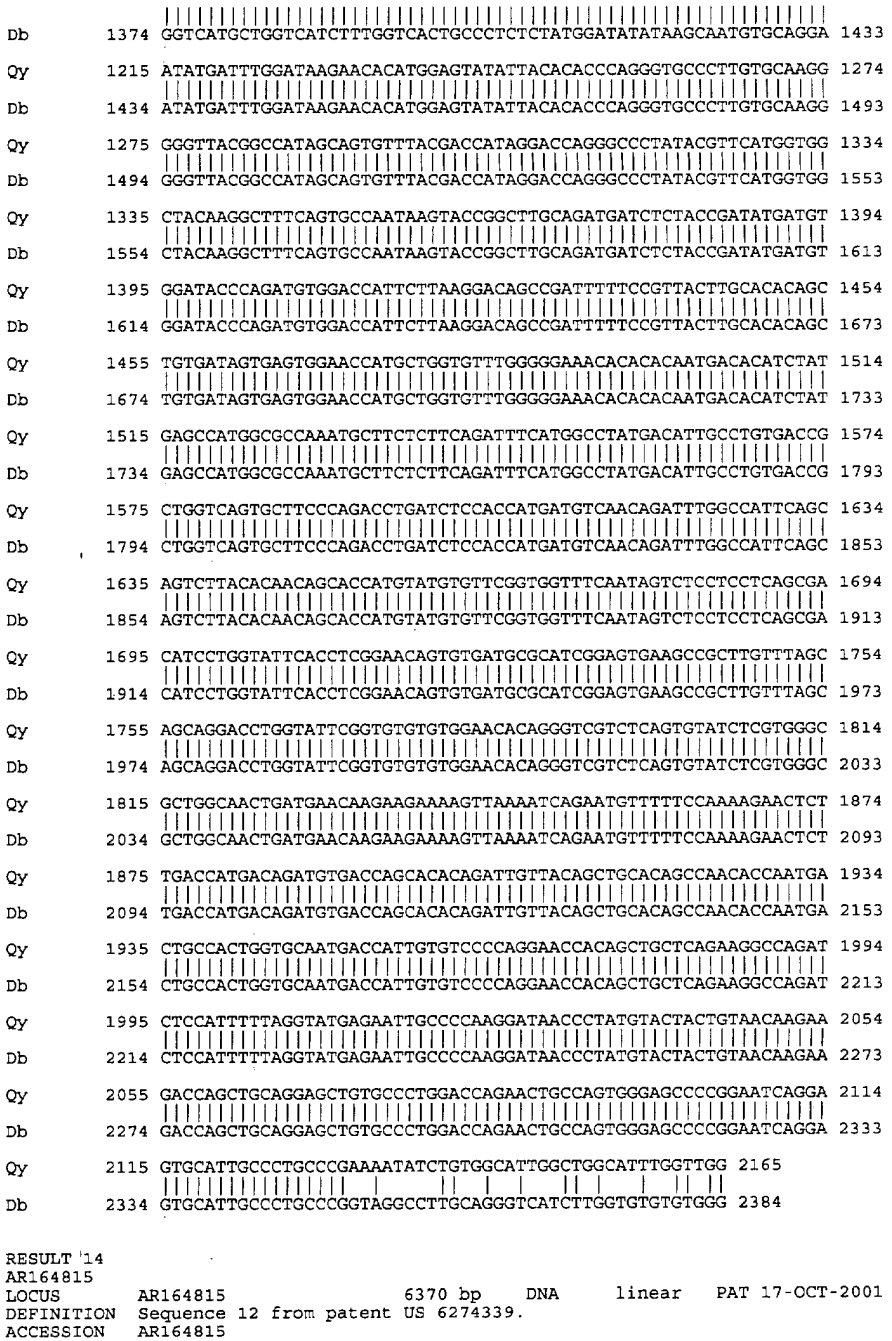
Qy       4068 A 4068
              |
Db       3181 A 3181
```

RESULT 13
AR164818
LOCUS AR164818 2625 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 18 from patent US 6274339.
ACCESSION AR164818
VERSION AR164818.1 GI:16238029
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 2625)
AUTHORS Moore,K. and Nagle,D.Lynn.
TITLE Methods and compositions for the diagnosis and treatment of body weight disorders, including obesity
JOURNAL Patent: US 6274339-A 18 14-AUG-2001;
FEATURES Location/Qualifiers
source 1. .2625
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN

Query Match 50.4%; Score 2051.8; DB 6; Length 2625;
Best Local Similarity 97.1%; Pred. No. 0;
Matches 2089; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

```
Qy         15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74
              || |||     |  ||   | ||  |  |  |  | | |   |||| ||   || || ||
Db        234 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAA 293
```





RESULT 14
AR164815
LOCUS AR164815 6370 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 12 from patent US 6274339.
ACCESSION AR164815

```
VERSION     AR164815.1  GI:16238024
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 6370)
  AUTHORS   Moore,K. and Nagle,D.Lynn.
  TITLE     Methods and compositions for the diagnosis and treatment of body
            weight disorders, including obesity
  JOURNAL   Patent: US 6274339-A 12 14-AUG-2001;
FEATURES             Location/Qualifiers
     source          1. .6370
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             45.6%;  Score 1853;  DB 6;  Length 6370;
  Best Local Similarity   96.3%;  Pred. No. 0;
  Matches 1980;  Conservative    7;  Mismatches   47;  Indels   23;  Gaps    8;

Qy       2035 ATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAG 2094
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAG 60

Qy       2095 TGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGG 2154
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGG 120

Qy       2155 CATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCT 2214
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCT 180

Qy       2215 AAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTA 2274
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTA 240

Qy       2275 GAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACC 2334
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACC 300

Qy       2335 TTAACCCCATGGGTCGGCCTT---CGGAAGATCAATGTGTCCTACTGG-----TGCTGGG 2386
             ||||||||||||||||| | |    | | | :|||||||||||||:|     |   |
Db        301 TTAACCCCATGGGTCGGGCCTTCGGGAAGGTYCAATGTGTCCTACTKGGTGCTGGGGAAG 360

Qy       2387 AAGATATGTCCCCATTTACAAATAG-TTTACTACAGTGGATGCC----GTCTGAGCCCAG 2441
              | || :  |||  |||||||||| |||||||||||    :     |   ||||||
Db        361 GATATGKTCCCCATTTTACAAATAGTTTTACTACAGTGGGATGSCCGTCTTGAGGCCCAG 420

Qy       2442 TGATGCT--GGATTCTGTGGAATTTT-----ATCAGAACCCAGTACTCGGGGACTGAAGG 2494
              || ||||  |:|||||||| | |||     |  |  |          |||||||||||||
Db        421 TGTTGCTTGGRATTCTGTGGGAATTTTATTCAGGAACCCAGTTACTTCGGGGACTGAAGG 480

Qy       2495 CTGCAACCTGCAT--CAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAG 2552
             ||||||||||||    | ||   ::|||||||||||||||||||||||||||||||||||
Db        481 CTGCAACCTGCATTCAACCCACTYMAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAG 540

Qy       2553 TGCTAA-GCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCG 2611
             |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGCTAAGGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCG 600

Qy       2612 GCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATG 2671
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATG 660

Qy       2672 TGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTG 2731
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTG 720

Qy       2732 AAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGT 2791
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGT 780

Qy       2792 GTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAG 2851
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAG 840

Qy       2852 TGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCA 2911
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCA 900

Qy       2912 GCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAAT 2971
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAAT 960

Qy       2972 GCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCA 3031
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCA 1020

Qy       3032 CAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGA 3091
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGA 1080

Qy       3092 AATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGT 3151
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGT 1140

Qy       3152 GCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATC 3211
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATC 1200

Qy       3212 GATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGT 3271
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGT 1260

Qy       3272 TCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTA 3331
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTA 1320

Qy       3332 CTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACC 3391
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACC 1380

Qy       3392 TCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTG 3451
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTG 1440

Qy       3452 TTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATT 3511
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATT 1500

Qy       3512 TTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCA 3571
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCA 1560

Qy       3572 AAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCTGGTACAGTTCTTCGTGA 3631
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCTGGTACAGTTCTTCGTGA 1620

Qy       3632 CTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGTGGTTTGGAAGATCAAAC 3691
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGTGGTTTGGAAGATCAAAC 1680

Qy       3692 AAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGAGATGCAACAGATGGCCA 3751
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGAGATGCAACAGATGGCCA 1740

Qy       3752 GCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTA 3811
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTA 1800

Qy       3812 TTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGAGCCGTGTTTTGGCAACA 3871
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 TTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGAGCCGTGTTTTGGCAACA 1860

Qy       3872 AAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCTGGGTGGCATCCCTCCTC 3931
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1861 AAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCTGGGTGGCATCCCTCCTC 1920

Qy       3932 CTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGA 3991
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 CTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGA 1980

Qy       3992 TAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGC 4051
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGC 2040

Qy       4052 CTGGGACCTGCATCTGA 4068
              ||||||||||||||||
Db       2041 CTGGGACCTGCATCTGA 2057


RESULT 15
AK000356
LOCUS       AK000356               2578 bp    mRNA    linear   PRI 13-SEP-2003
DEFINITION  Homo sapiens cDNA FLJ20349 fis, clone HEP13957, highly similar to
            AB011120 Homo sapiens mRNA for KIAA0548 protein.
ACCESSION   AK000356
VERSION     AK000356.1  GI:7020387
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T.,
            Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Ota,T.,
            Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
            Nakamura,Y., Isogai,T. and Sugano,S.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2578)
  AUTHORS   Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
            Shibahara,T., Tanaka,T. and Nakamura,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science,
            University of Tokyo, Deptment of Virology; Shirokane-dai, 4-6-1,
            Minato-ku, Tokyo 108-8639, Japan (E-mail:flcdna@ims.u-tokyo.ac.jp,
            Tel:81-3-5449-5286, Fax:81-3-5449-5416)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            International Trade and Industry of Japan; cDNA full insert
            sequencing: Research Association for Biotechnology; cDNA library
            construction, 5'- & 3'-end one pass sequencing: Departent of
            Virology and Human Genome Center, Institute of Medical Science,
            University of Tokyo (partly supported by Science and Technology
            Agency).
FEATURES             Location/Qualifiers
     source          1. .2578
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HEP13957"
                     /cell_line="HepG2"
                     /cell_type="hepatoma"
                     /clone_lib="HEP"
                     /note="cloning vector pME18SFL3"
     misc_feature    1. .2578
                     /note="highly similar to AB011120 Homo sapiens mRNA for
                     KIAA0548 protein"
ORIGIN

  Query Match           41.7%;  Score 1697;  DB 9;  Length 2578;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1697;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2372 CCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGT 2431
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGT 60

Qy       2432 CTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGA 2491
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGA 120

Qy       2492 AGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACA 2551
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACA 180

Qy       2552 GTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCG 2611
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCG 240

Qy       2612 GCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATG 2671
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATG 300

Qy       2672 TGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTG 2731
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTG 360

Qy       2732 AAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGT 2791
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGT 420

Qy       2792 GTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAG 2851
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAG 480

Qy       2852 TGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCA 2911
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCA 540

Qy       2912 GCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAAT 2971
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAAT 600

Qy       2972 GCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCA 3031
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCA 660

Qy       3032 CAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGA 3091
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGA 720

Qy       3092 AATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGT 3151
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGT 780

Qy       3152 GCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATC 3211
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATC 840

Qy       3212 GATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGT 3271
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGT 900

Qy       3272 TCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTA 3331
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTA 960

Qy       3332 CTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACC 3391
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACC 1020

Qy       3392 TCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTG 3451
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTG 1080

Qy       3452 TTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATT 3511
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATT 1140

Qy       3512 TTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCA 3571
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1141 TTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCA 1200

Qy       3572 AAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCTGGTACAGTTCTTCGTGA 3631
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCTGGTACAGTTCTTCGTGA 1260

Qy       3632 CTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGTGGTTTGGAAGATCAAAC 3691
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGTGGTTTGGAAGATCAAAC 1320

Qy       3692 AAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGAGATGCAACAGATGGCCA 3751
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 AAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGAGATGCAACAGATGGCCA 1380

Qy       3752 GCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTA 3811
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTA 1440

Qy       3812 TTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGAGCCGTGTTTTGGCAACA 3871
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGAGCCGTGTTTTGGCAACA 1500

Qy       3872 AAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCTGGGTGGCATCCCTCCTC 3931
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 AAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCTGGGTGGCATCCCTCCTC 1560

Qy       3932 CTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGA 3991
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGA 1620

Qy       3992 TAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGC 4051
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGC 1680

Qy       4052 CTGGGACCTGCATCTGA 4068
             ||||||||||||||||
Db       1681 CTGGGACCTGCATCTGA 1697
```

Search completed: March 5, 2004, 05:11:12
Job time : 10247.8 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2004, 16:59:34 ; Search time 6445.51 Seconds
(without alignments)
18847.119 Million cell updates/sec

Title: US-09-787-097-11
Perfect score: 4068
Sequence: 1 atggtggccgcagcggcggc.....agcctgggacctgcatctga 4068

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

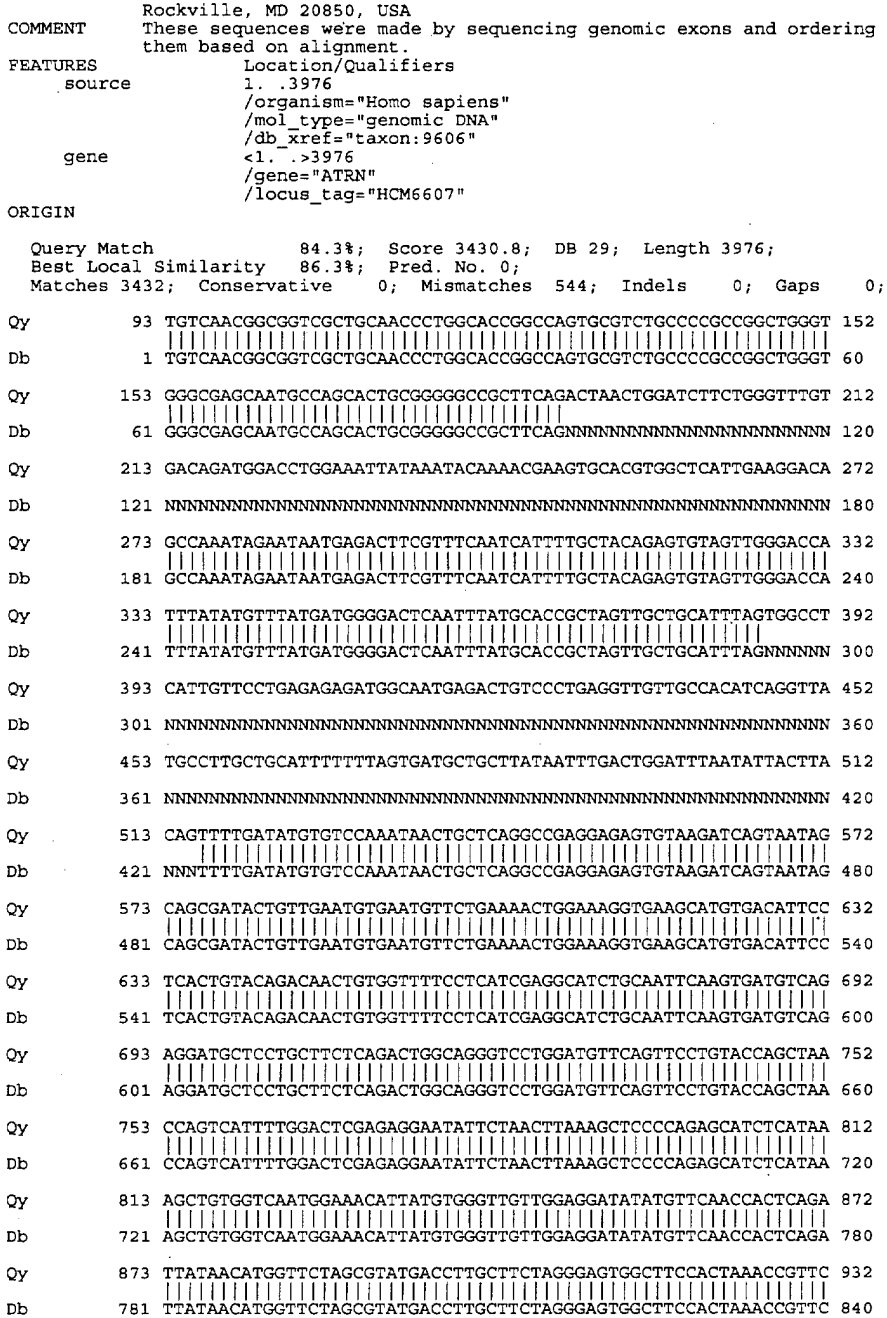
Result No.	Score	% Query Match	Length	DB	ID	Description
1	3430.8	84.3	3976	29	AY418587	AY418587 Homo sapi
2	2923.6	71.9	3976	29	AY418589	AY418589 Mus muscu
3	2866	70.5	3671	29	AY418588	AY418588 Pan trogl
4	925.6	22.8	1201	13	BX440935	BX440935 BX440935
5	775	19.1	964	13	BU506373	BU506373 AGENCOURT
6	758.8	18.7	933	12	BG678679	BG678679 602624477
7	699.6	17.2	840	14	CB520374	CB520374 UI-M-GI0-
8	681	16.7	681	12	BM783739	BM783739 K-EST0061
9	680.4	16.7	683	12	BM783788	BM783788 K-EST0061
10	680	16.7	680	12	BM783258	BM783258 K-EST0061
11	680	16.7	680	12	BM783773	BM783773 K-EST0061
12	677	16.6	677	12	BM783718	BM783718 K-EST0061
13	674.4	16.6	677	12	BM783821	BM783821 K-EST0061
14	673	16.5	673	12	BM783685	BM783685 K-EST0061
15	665.4	16.4	792	14	CD644584	CD644584 AGENCOURT
16	652.2	16.0	962	13	BX431593	BX431593 BX431593
17	636.8	15.7	674	12	BG563729	BG563729 602584512
18	636.4	15.6	770	13	BY761276	BY761276 BY761276
19	636	15.6	777	14	CF745736	CF745736 UI-M-GV0-
20	634	15.6	706	12	BI089332	BI089332 602853292
21	632	15.5	632	12	BM773003	BM773003 K-EST0057
22	632	15.5	632	12	BM783264	BM783264 K-EST0061
23	628.2	15.4	953	10	BF313142	BF313142 601897661
24	623	15.3	917	10	BF672370	BF672370 602150605
25	618	15.2	633	10	BE894518	BE894518 601433126
c 26	612.2	15.0	722	13	BX669806	BX669806 BX669806
27	608	14.9	735	13	BY761390	BY761390 BY761390
28	607.8	14.9	991	10	BE880905	BE880905 601490448
29	599.2	14.7	940	13	BQ956223	BQ956223 AGENCOURT
30	589.6	14.5	743	14	CB519923	CB519923 UI-M-GI0-
31	585.2	14.4	716	14	CF745901	CF745901 UI-M-GV0-
c 32	580.2	14.3	714	10	BE650220	BE650220 UI-M-BH2.
33	580	14.3	591	12	BI850289	BI850289 imageqc_1
34	569.8	14.0	653	14	CB446934	CB446934 700912 MA
35	554.2	13.6	663	14	CF745835	CF745835 UI-M-GV0-
36	549.6	13.5	902	13	BU151781	BU151781 AGENCOURT
c 37	539	13.2	644	14	CB425680	CB425680 600762 MA
38	537.4	13.2	942	10	BF316716	BF316716 601903293
39	532.2	13.1	641	10	BB664616	BB664616 BB664616
40	524.2	12.9	626	10	BE980188	BE980188 UI-M-BG2-
41	522.6	12.8	600	14	CB432517	CB432517 608730 MA
42	517.6	12.7	595	10	BF853681	BF853681 MR2-EN009
43	516.6	12.7	598	10	BF853006	BF853006 MR2-EN009
44	513.2	12.6	544	10	BF853962	BF853962 MR2-EN009
c 45	502.6	12.4	524	10	BE709733	BE709733 IL3-HT061

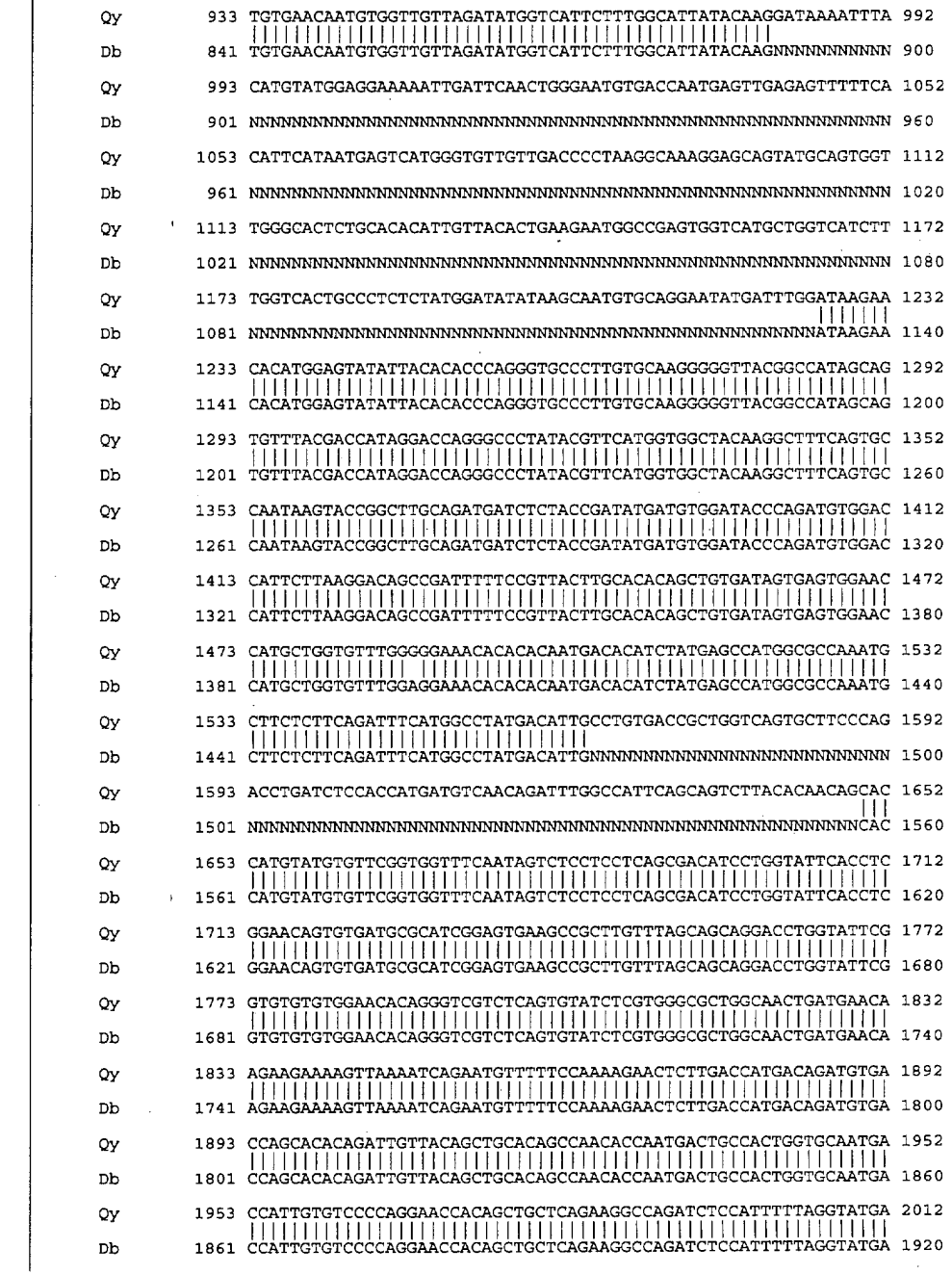
ALIGNMENTS

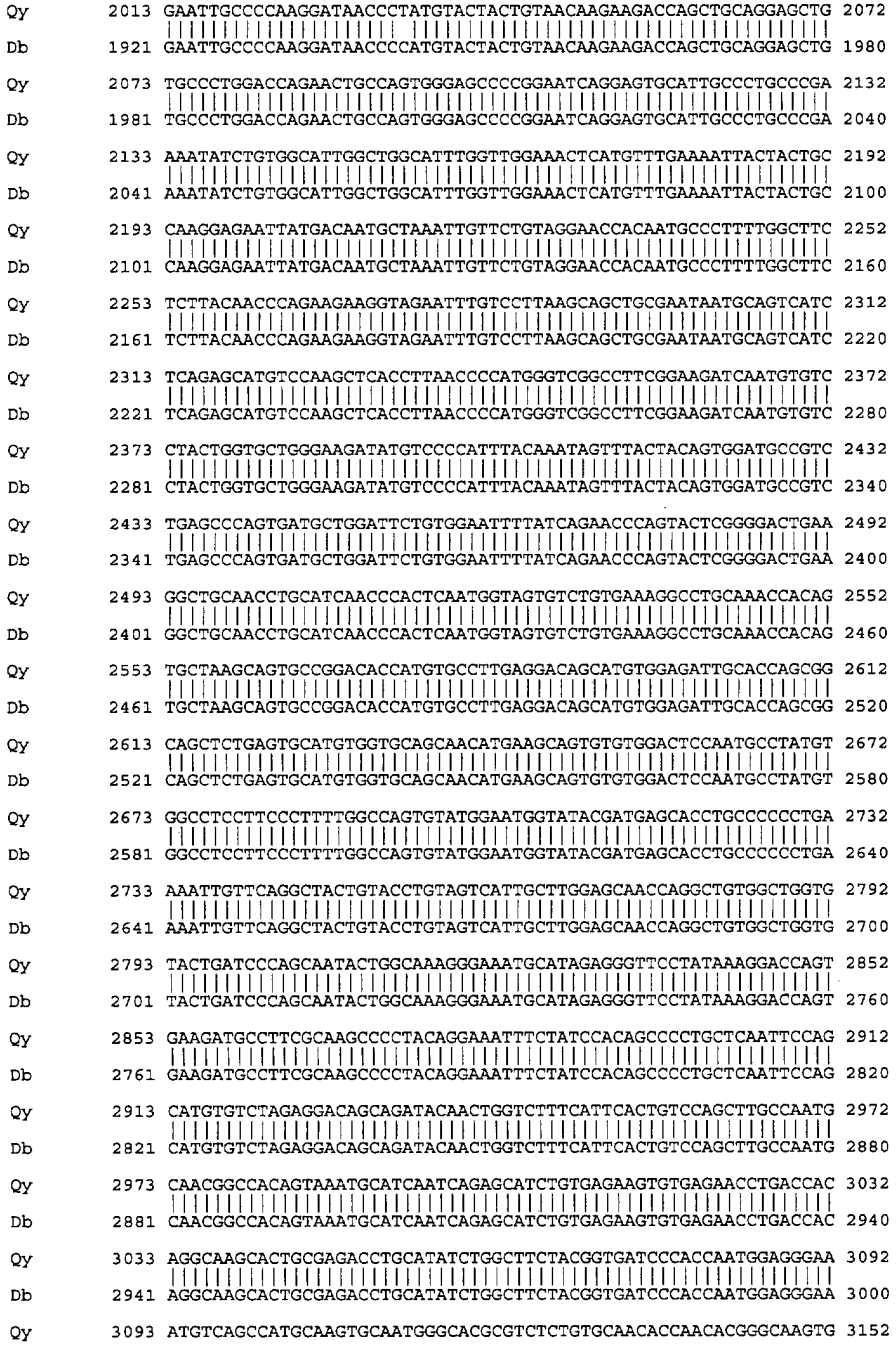
RESULT 1
AY418587
LOCUS AY418587 3976 bp DNA linear GSS 17-DEC-2003
DEFINITION Homo sapiens ATRN gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION AY418587
VERSION AY418587.1 GI:39774547
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3976)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 3976)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission

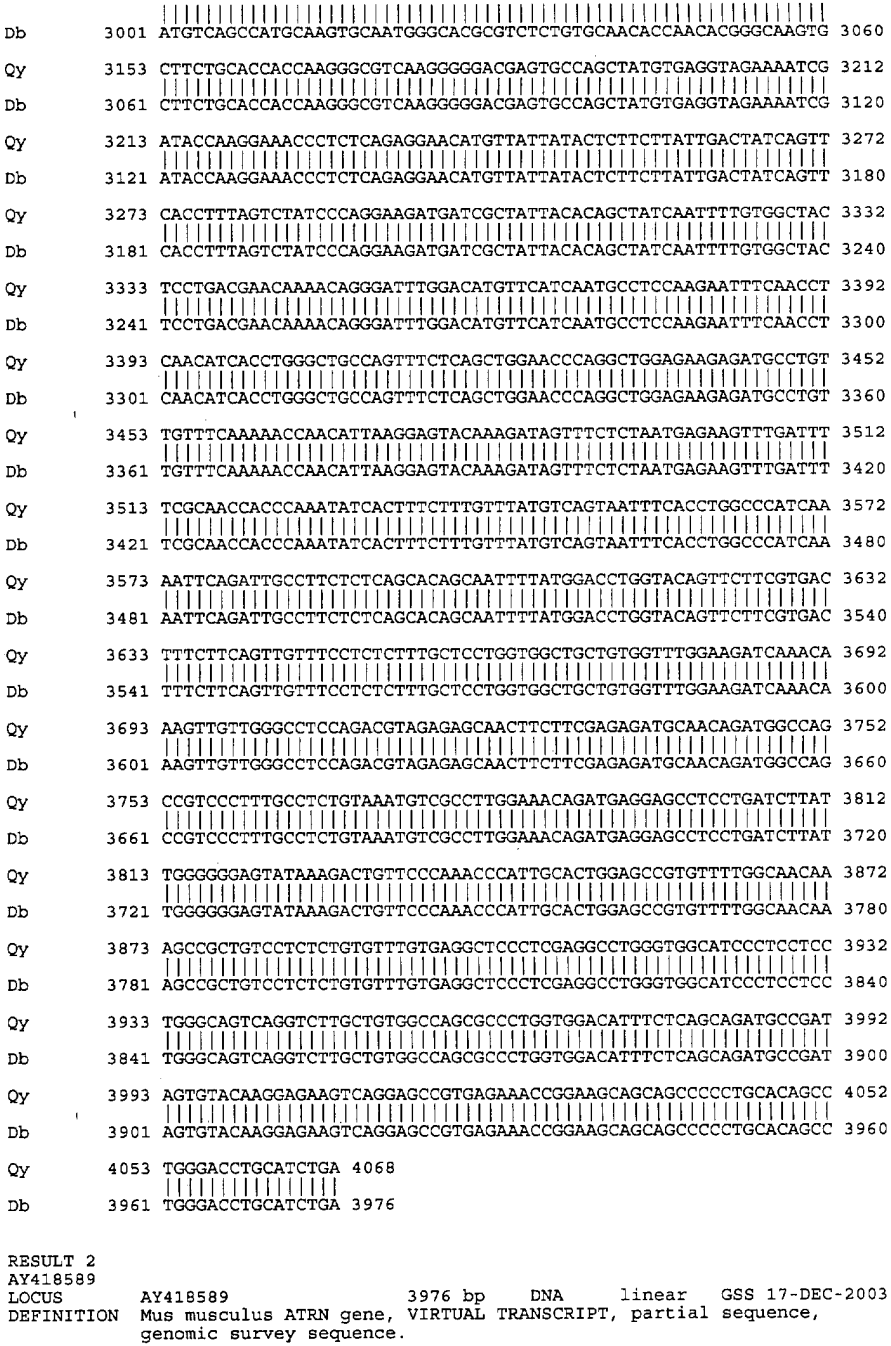
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT These sequences were made by sequencing genomic exons and ordering them based on alignment.
FEATURES Location/Qualifiers
source 1. .3976
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
gene <1. .>3976
/gene="ATRN"
/locus_tag="HCM6607"
ORIGIN

Query Match 84.3%; Score 3430.8; DB 29; Length 3976;
Best Local Similarity 86.3%; Pred. No. 0;
Matches 3432; Conservative 0; Mismatches 544; Indels 0; Gaps 0;









RESULT 2
AY418589
LOCUS AY418589 3976 bp DNA linear GSS 17-DEC-2003
DEFINITION Mus musculus ATRN gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.

```
ACCESSION   AY418589
VERSION     AY418589.1  GI:39774549
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 3976)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 3976)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     These sequences were made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .3976
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
     gene            <1. .>3976
                     /gene="ATRN"
                     /locus_tag="HCM6607"
ORIGIN

  Query Match             71.9%;  Score 2923.6;  DB 29;  Length 3976;
  Best Local Similarity   78.3%;  Pred. No. 0;
  Matches 3115;  Conservative    0;  Mismatches  861;  Indels    0;  Gaps    0;

Qy          93 TGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGT 152
               |||||||||||| |||||||||||||||||||||||||||||||||| | ||||||||
Db           1 TGTCAACGGCGGCCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCACGGGCTGGGT 60

Qy         153 GGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGT 212
               ||||||||||||||||||||||||||||||||||
Db          61 GGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGNNNNNNNNNNNNNNNNNNNNNNN 120

Qy         213 GACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGACA 272

Db         121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 180

Qy         273 GCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCA 332
               |||||||||||||||||||||||||| ||||| |||||||||||| ||||| ||||||||
Db         181 GCCAAATAGAATAATGAGACTTCGCTTCAACCATTTTGCTACAGAATGTAGCTGGGACCA 240

Qy         333 TTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCT 392
               |||||||||||||||||||||||||| || ||||| ||  ||||||| |||||
Db         241 TTTATATGTTTATGATGGGGACTCAATCTACGCACCTCTGATTGCTGCCTTTAGNNNNNN 300

Qy         393 CATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTA 452

Db         301 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 360

Qy         453 TGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTA 512

Db         361 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 420

Qy         513 CAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAG 572
               |||||| |||||||| ||||| ||||||||||||||||||||||||| |||||| ||
Db         421 NNNTTTTGACATGTGTCCGAATAATTGCTCAGGCCGAGGAGAGTGTAAGAGCAGTAACAG 480

Qy         573 CAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCC 632

               ||||   |||||||| |||||||||||||||||||||||||||| ||  | ||||||||||
Db         481 CAGCAGCGCTGTTGAGTGTGAATGTTCTGAAAACTGGAAAGGGGAGTCGTGTGACATTCC 540

Qy         633 TCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAG 692
               |||||||||||||||||||||| |||||||| |||||||||| ||| |||| |||  |||
Db         541 TCACTGTACAGACAACTGTGGCTTTCCTCACCGAGGCATCTGTAATGCAAGCGATACCAG 600

Qy         693 AGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAA 752
               ||| |||||||||||   |||||||||||||||||||| ||| |||||||  ||||||||
Db         601 AGGGTGCTCCTGCTTTCCTCACTGGCAGGGTCCTGGATGCTCAATTCCTGTGCCAGCTAA 660

Qy         753 CCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAA 812
               |||||| |||||||||||||| |||||||| | |||||||| |||||||  ||||||||
Db         661 CCAGTCTTTTTGGACTCGAGAAGAATATTCTGATTTAAAGCTTCCCAGAGCCTCTCATAA 720

Qy         813 AGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGA 872
               |||||||||||||||||  || |||||||||||||| |||||||||||||||||  |||||
Db         721 AGCTGTGGTCAATGGAAATATAATGTGGGTTGTTGGCGGATATATGTTCAACCATTCAGA 780

Qy         873 TTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTC 932
               ||| | |||||||| |||||||||||||  |||||||||  |||||||||||||||| |||
Db         781 TTACAGCATGGTTTTAGCGTATGACCTGACTTCTAGGGAATGGCTTCCACTAAACCATTC 840

Qy         933 TGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTA 992
               |||||||| |||||||| |||||||||||||||||||||||| | |||
Db         841 TGTGAACAGTGTGGTTGTAAGATATGGTCATTCTTTGGCATTACATAAGNNNNNNNNNNN 900

Qy         993 CATGTATGGAGGAAAAAATTGATTCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCA 1052

Db         901 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 960

Qy        1053 CATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGT 1112

Db         961 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1020

Qy        1113 TGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTT 1172

Db        1021 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1080

Qy        1173 TGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAA 1232
                                                                | |||||
Db        1081 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNAAAAGAA 1140

Qy        1233 CACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAG 1292
               |||||||||||||||||  || |||||||| |||||||||||||||||  ||||| || ||
Db        1141 CACATGGAGTATATTACATACTCAGGGTGCTCTTGTGCAAGGGGGTTATGGCCACAGTAG 1200

Qy        1293 TGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGC 1352
               |||||| ||  | ||||||| ||| || ||||||||||||||||||||||||||||| ||
Db        1201 TGTTTATGATGACAGGACCAAGGCTCTGTACGTTCATGGTGGCTACAAGGCTTTCAGCGC 1260

Qy        1353 CAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGAC 1412
               ||| || |||||||||||||||||| |||||| |||| |||||||||| ||||||||||
Db        1261 CAACAAATACCGGCTTGCAGATGACCTCTACAGATACGATGTGGATACTCAGATGTGGAC 1320

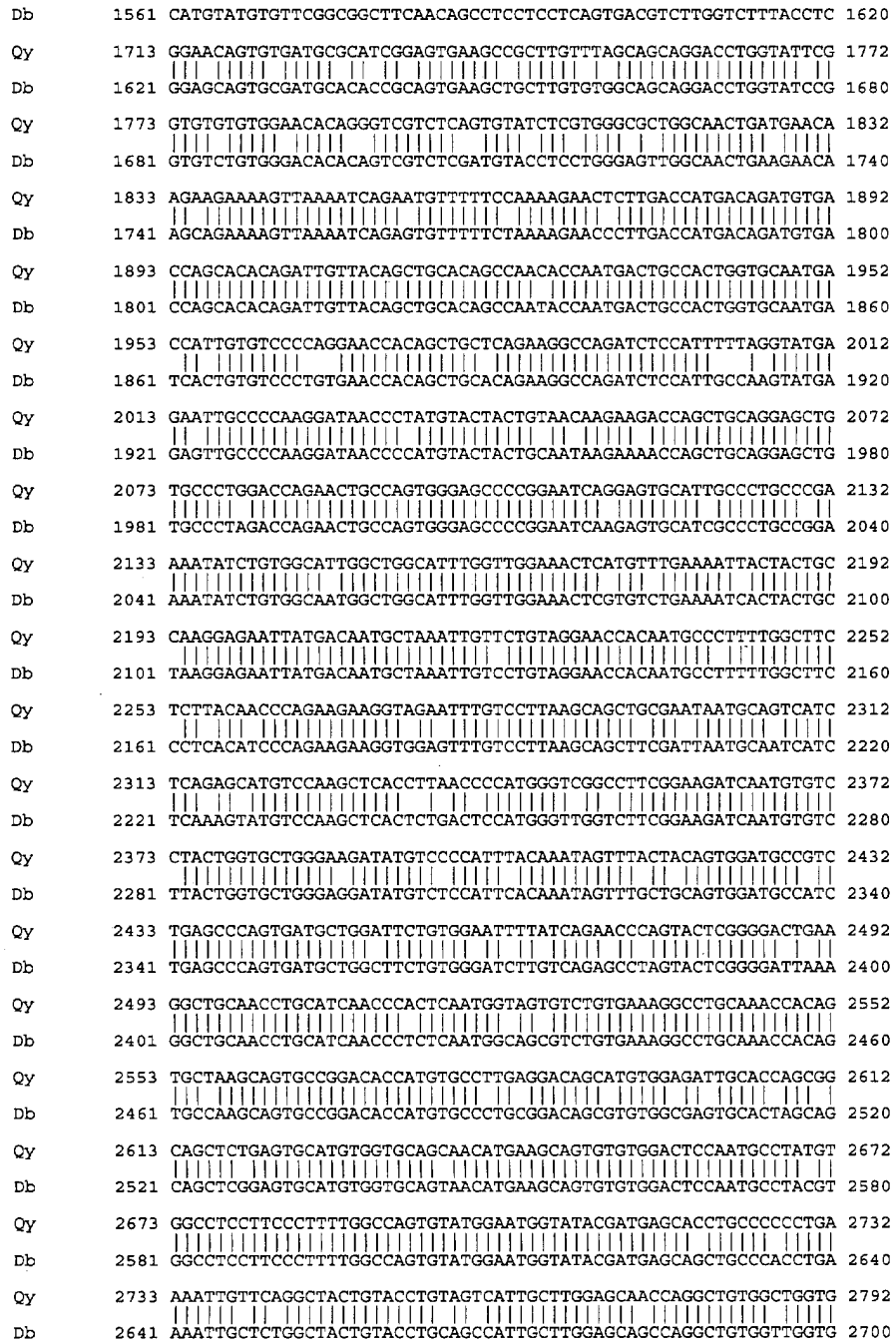
Qy        1413 CATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAAC 1472
               |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        1321 CATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCATACAGCTGTGATAGTGAGTGGAAC 1380

Qy        1473 CATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATG 1532
               |||||||||||||||  || |||||||||||||||  |||||||| || |||||||||
Db        1381 CATGCTGGTGTTTGGAGGGAACACACACAATGACACTTCCATGAGCCACGGTGCCAAATG 1440

Qy        1533 CTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAG 1592
               |||||| |||| |||||||| ||||||||
Db        1441 CTTCTCCTCAGACTTCATGGCTTATGACATTGNNNNNNNNNNNNNNNNNNNNNNNNNNN 1500

Qy        1593 ACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACAACAGCAC 1652
                                                                    |||
Db        1501 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCAC 1560

Qy        1653 CATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTC 1712
               |||||||||||||| || ||||| || |||||||||| ||| || |||| || |||||
```


Db 1561 CATGTATGTGTTCGGCGGCTTCAACAGCCTCCTCCTCAGTGACGTCTTGGTCTTTACCTC 1620
Qy 1713 GGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCG 1772
Db 1621 GGAGCAGTGCGATGCACACCGCAGTGAAGCTGCTTGTGTGGCAGCAGGACCTGGTATCCG 1680
Qy 1773 GTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACA 1832
Db 1681 GTGTCTGTGGGACACACAGTCGTCTCGATGTACCTCCTGGGAGTTGGCAACTGAAGAACA 1740
Qy 1833 AGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGA 1892
Db 1741 AGCAGAAAAGTTAAAATCAGAGTGTTTTTCTAAAAGAACCCTTGACCATGACAGATGTGA 1800
Qy 1893 CCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGCAATGA 1952
Db 1801 CCAGCACACAGATTGTTACAGCTGCACAGCCAATACCAATGACTGCCACTGGTGCAATGA 1860
Qy 1953 CCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGA 2012
Db 1861 TCACTGTGTCCCTGTGAACCACAGCTGCACAGAAGGCCAGATCTCCATTGCCAAGTATGA 1920
Qy 2013 GAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTG 2072
Db 1921 GAGTTGCCCCAAGGATAACCCCATGTACTACTGCAATAAGAAAACCAGCTGCAGGAGCTG 1980
Qy 2073 TGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGA 2132
Db 1981 TGCCCTAGACCAGAACTGCCAGTGGGAGCCCCGGAATCAAGAGTGCATCGCCCTGCCGGA 2040
Qy 2133 AAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGC 2192
Db 2041 AAATATCTGTGGCAATGGCTGGCATTTGGTTGGAAACTCGTGTCTGAAAATCACTACTGC 2100
Qy 2193 CAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTC 2252
Db 2101 TAAGGAGAATTATGACAATGCTAAATTGTCCTGTAGGAACCACAATGCCTTTTTGGCTTC 2160
Qy 2253 TCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATC 2312
Db 2161 CCTCACATCCCAGAAGAAGGTGGAGTTTGTCCTTAAGCAGCTTCGATTAATGCAATCATC 2220
Qy 2313 TCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTC 2372
Db 2221 TCAAAGTATGTCCAAGCTCACTCTGACTCCATGGGTTGGTCTTCGGAAGATCAATGTGTC 2280
Qy 2373 CTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTC 2432
Db 2281 TTACTGGTGCTGGGAGGATATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATC 2340
Qy 2433 TGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAA 2492
Db 2341 TGAGCCCAGTGATGCTGGCTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGATTAAA 2400
Qy 2493 GGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAG 2552
Db 2401 GGCTGCAACCTGCATCAACCCTCTCAATGGCAGCGTCTGTGAAAGGCCTGCAAACCACAG 2460
Qy 2553 TGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGG 2612
Db 2461 TGCCAAGCAGTGCCGGACACCATGTGCCCTGCGGACAGCGTGTGGCGAGTGCACTAGCAG 2520
Qy 2613 CAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGT 2672
Db 2521 CAGCTCGGAGTGCATGTGGTGCAGTAACATGAAGCAGTGTGTGGACTCCAATGCCTACGT 2580
Qy 2673 GGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGA 2732
Db 2581 GGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCAGCTGCCCACCTGA 2640
Qy 2733 AAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTG 2792
Db 2641 AAATTGCTCTGGCTACTGTACCTGCAGCCATTGCTTGGAGCAGCCAGGCTGTGGTTGGTG 2700
Qy 2793 TACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGT 2852
Db 2701 TACTGATCCTAGCAATACTGGGAAAGGAAAATGTATTGAGGGCAGCTATAAAGGACCTGT 2760
Qy 2853 GAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAG 2912
Db 2761 GAAGATGCCGTCACAGGCCTCTGCAGGAAATGTGTATCCACAGCCCCTTCTGAACTCCAG 2820
Qy 2913 CATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATG 2972
Db 2821 CATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAGTG 2880
Qy 2973 CAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCAC 3032
Db 2881 CAACGGACACAGCAAATGCATCAACCAGAGTATCTGTGAGAAGTGTGAGGACCTGACCAC 2940
Qy 3033 AGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAA 3092
Db 2941 GGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTATGGTGACCCGACTAATGGAGGCAA 3000
Qy 3093 ATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTG 3152
Db 3001 ATGTCAGCCATGCAAGTGCAATGGGCACGCATCACTGTGCAACACCAACACCGGCAAGTG 3060
Qy 3153 CTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCG 3212
Db 3061 CTTCTGTACCACCAAAGGTGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCG 3120
Qy 3213 ATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTT 3272
Db 3121 ATACCAGGGAAACCCTCTCAAAGGAACATGCTACTATACCCTTCTCATTGACTATCAGTT 3180
Qy 3273 CACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTAC 3332
Db 3181 CACCTTTAGCCTGTCCCAGGAAGACGACCGCTACTACACAGCCATCAACTTTGTGGCTAC 3240
Qy 3333 TCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCT 3392
Db 3241 TCCTGATGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAAAACTTCAACCT 3300
Qy 3393 CAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGT 3452
Db 3301 CAACATCACCTGGGCCACCAGCTTCCCAGCCGGAACCCAGACTGGAGAAGAGGTGCCTGT 3360
Qy 3453 TGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTT 3512
Db 3361 TGTTTCAAAAACCAACATCAAGGAATACAAAGATAGCTTCTCTAATGAGAAATTTGATTT 3420
Qy 3513 TCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAA 3572
Db 3421 TCGCAACCATCCAAACATCACTTTCTTTGTTTATGTCAGTAATTTCACTTGGCCCATCAA 3480
Qy 3573 AATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCTGGTACAGTTCTTCGTGAC 3632
Db 3481 AATTCAGATTGCCTTCTCCCAGCACAGCAACTTCATGGACCTGGTACAGTTCTTCGTGAC 3540
Qy 3633 TTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGTGGTTTGGAAGATCAAACA 3692
Db 3541 TTTCTTCAGTTGTTTCCTCTCGCTGCTTCTGGTGGCTGCAGTGGTCTGGAAGATCAAGCA 3600
Qy 3693 AAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGAGATGCAACAGATGGCCAG 3752
Db 3601 GAGCTGTTGGGCATCCAGGCGGAGAGAGCAACTTCTTCGGGAGATGCAACAGATGGCCAG 3660
Qy 3753 CCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTAT 3812
Db 3661 CCGCCCCTTTGCTTCTGTAAACGTTGCCTTGGAAACAGATGAAGAACCTCCTGATCTCAT 3720
Qy 3813 TGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGAGCCGTGTTTTGGCAACAA 3872
Db 3721 TGGGGGAAGTATAAAGACCGTTCCTAAGCCCATTGCCCTGGAGCCCTGCTTTGGTAACAA 3780

```
Qy       3873 AGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCTGGGTGGCATCCCTCCTCC 3932
             |||||| |||||||||| || ||||||||||||||| ||||| || ||||||||||
Db       3781 AGCCGCAGTCCTCTCTGTATTCGTGAGGCTCCCTCGAGGACTGGGAGGAATCCCTCCTCC 3840

Qy       3933 TGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGAT 3992
             ||| |||||||||||| |||||||||| |||||||||||||||||||||||||||| ||
Db       3841 TGGTCAGTCAGGTCTCGCTGTGGCCAGTGCCCTGGTGGACATTTCTCAGCAGATGCCAAT 3900

Qy       3993 AGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGCC 4052
             |||||||||||||||||||||||||| || ||||||||||||||||| ||||||||||
Db       3901 AGTGTACAAGGAGAAGTCAGGAGCTGTAAGAAACCGGAAGCAGCAGCCGCCTGCACAGCC 3960

Qy       4053 TGGGACCTGCATCTGA 4068
             ||| |||||||| |||
Db       3961 TGGAACCTGCATTTGA 3976
```

RESULT 3
AY418588
LOCUS AY418588 3671 bp DNA linear GSS 17-DEC-2003
DEFINITION Pan troglodytes ATRN gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION AY418588
VERSION AY418588.1 GI:39774548
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 3671)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 3671)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT These sequences were made by sequencing genomic exons and ordering them based on alignment.
FEATURES Location/Qualifiers
source 1. .3671
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
gene <1. .>3671
/gene="ATRN"
/locus_tag="HCM6607"
ORIGIN

Query Match 70.5%; Score 2866; DB 29; Length 3671;
Best Local Similarity 78.3%; Pred. No. 0;
Matches 2875; Conservative 0; Mismatches 796; Indels 0; Gaps 0;

```
Qy        273 GCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCA 332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCA 60

Qy        333 TTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCT 392
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGNNNNNN 120

Qy        393 CATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTA 452
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 180

Qy        453 TGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTA 512
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 240

Qy        513 CAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAG 572
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NNNTTTTGGTATGTGTCCGAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAG 300

Qy        573 CAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCC 632
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCC 360

Qy        633 TCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAG 692
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAG 420

Qy        693 AGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAA 752
              |||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||
Db        421 AGGATGCTCCTGCTTCTCAGACTGGCAGGNNNCTGGATGTTCAGTTCCTGTACCAGCTAA 480

Qy        753 CCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAA 812
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAA 540

Qy        813 AGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGA 872
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGA 600

Qy        873 TTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTC 932
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTC 660

Qy        933 TGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTA 992
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATACAAGNNNNNNNNNN 720

Qy        993 CATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCA 1052
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 780

Qy       1053 CATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGT 1112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 840

Qy       1113 TGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTT 1172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 900

Qy       1173 TGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAA 1232
                                                                  |||||||
Db        901 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNATAAGAA 960

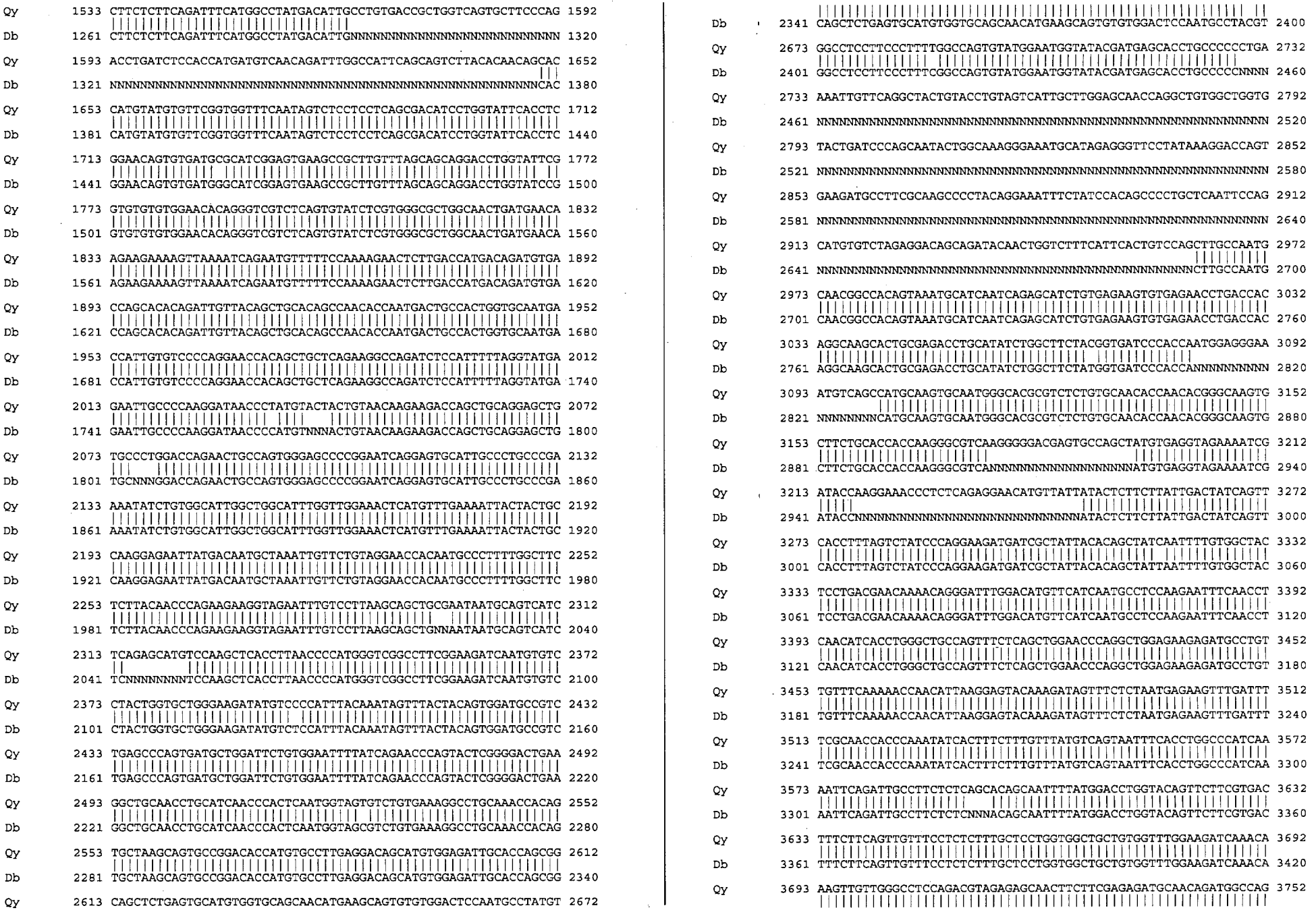
Qy       1233 CACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAG 1292
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CACGTGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAG 1020

Qy       1293 TGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGC 1352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGC 1080

Qy       1353 CAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGAC 1412
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGAC 1140

Qy       1413 CATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAAC 1472
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAAC 1200

Qy       1473 CATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATG 1532
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CATGCTGGTGTTTGGAGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATG 1260
```



```
Db       3421 AAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGAGATGCAACAGATGGCCAG 3480

Qy       3753 CCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTAT 3812
             |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       3481 CCGTCCCTTTGCCTCCGTAAATGTCGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTAT 3540

Qy       3813 TGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGAGCCGTGTTTTGGCAACAA 3872
             |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db       3541 TGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGAGCCATGTTTTGGCAACAA 3600

Qy       3873 AGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCTGGGTGGCATCCCTCCTCC 3932
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3601 AGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCTGGGTGGCATCCCTCCTCC 3660

Qy       3933 TGGGCAGTCAG 3943
             ||||||||||||
Db       3661 TGGGCAGTCAG 3671


RESULT 4
BX440935
LOCUS       BX440935               1201 bp    mRNA    linear   EST 15-MAY-2003
DEFINITION  BX440935 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
            CS0DF012YL19 5-PRIME, mRNA sequence.
ACCESSION   BX440935
VERSION     BX440935.1  GI:30781889
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 10212.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DF012CF10QP1&cluster=10212.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DF012CF10QP1.
FEATURES             Location/Qualifiers
     source          1. .1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DF012YL19"
                     /tissue_type="FETAL BRAIN"
                     /dev_stage="fetal"
                     /clone_lib="Homo sapiens FETAL BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN

  Query Match             22.8%;  Score 925.6;  DB 13;  Length 1201;
  Best Local Similarity   98.4%;  Pred. No. 3.4e-224;
  Matches  962;  Conservative    4;  Mismatches    8;  Indels    4;  Gaps    3;

Qy       2115 GTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATG 2174
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         57 GTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATG 116

Qy       2175 TTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCA 2234
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        117 TTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCA 176

Qy       2235 CAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCT 2294
             |||||||||||||||||||||||||||||| || ||||||||||||||||||||||||||
Db        177 CAATGCCCTTTTGGCTTCTCTTACAACCCTGATGAAGGTAGAATTTGTCCTTAAGCAGCT 236

Qy       2295 GCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCT 2354
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        237 GCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCT 296

Qy       2355 TCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTT 2414
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 TCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTT 356

Qy       2415 ACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACC 2474
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 ACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACC 416

Qy       2475 CAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGA 2534
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        417 CAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGA 476

Qy       2535 AAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATG 2594
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        477 AAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATG 536

Qy       2595 TGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGT 2654
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        537 TGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGT 596

Qy       2655 GGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGAT 2714
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 GGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGAT 656

Qy       2715 GAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCA 2774
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 GAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCA 716

Qy       2775 ACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGG 2834
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 ACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGG 776

Qy       2835 TTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACA 2894
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 TTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACA 836

Qy       2895 GCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCA 2954
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 GCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCA 896

Qy       2955 CTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAA 3014
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 CTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAA 956

Qy       3015 GTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGA 3074
             |||||||||:  ||   : | :|||||||||||| |||||||||||| |||||||:||
Db        957 GTGTGAGAACYGAAC--AMAGGMAGCACTGCGAGAACTGCATATCTGGC-TCTACGGKGA 1013

Qy       3075 TCCCACCAATGGAGGGAA 3092
             | ||||||||||||||||
Db       1014 T-CCACCAATGGAGGGAA 1030


RESULT 5
BU506373
LOCUS       BU506373               964 bp    mRNA    linear   EST 12-SEP-2002
DEFINITION  AGENCOURT_10015485 NIH_MGC_94 Mus musculus cDNA clone IMAGE:6494441
            5', mRNA sequence.
ACCESSION   BU506373
VERSION     BU506373.1  GI:22812606
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 964)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: The Cepko Laboratory
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM14050  row: k  column: 18
            High quality sequence stop: 592.
FEATURES             Location/Qualifiers
     source          1. .964
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:6494441"
                     /tissue_type="retina"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_94"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 3.3 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             19.1%;  Score 775;  DB 13;  Length 964;
  Best Local Similarity   88.9%;  Pred. No. 7.8e-186;
  Matches  857;  Conservative    0;  Mismatches  105;  Indels    2;  Gaps    2;

Qy       1774 TGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAA 1833
Db          1 TGTCTGTGGGACACACAGTCGTCTCGATGTACCTCCTGGGAGTTGGCAACTGAAGAACAA 60

Qy       1834 GAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGAC 1893
Db         61 GCAGAAAAGTTAAAATCAGAGTGTTTTTCTAAAAGAACCCTTGACCATGACAGATGTGAC 120

Qy       1894 CAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGCAATGAC 1953
Db        121 CAGCACACAGATTGTTACAGCTGCACAGCCAATACCAATGACTGCCACTGGTGCAATGAT 180

Qy       1954 CATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAG 2013
Db        181 CACTGTGTCCCTGTGAACCACAGCTGCACAGAAGGCCAGATCTCCATTGCCAAGTATGAG 240

Qy       2014 AATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGT 2073
Db        241 AGTTGCCCCAAGGATAACCCCATGTACTACTGCAATAAGAAAACCAGCTGCAGGAGCTGT 300

Qy       2074 GCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAA 2133
Db        301 GCCCTAGACCAGAACTGCCAGTGGGAGCCCCGGAATCAAGAGTGCATCGCCCTGCCGGAA 360

Qy       2134 AATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCC 2193
Db        361 AATATCTGTGGCAATGGCTGGCATTTGGTTGGAAACTCGTGTCTGAAAATCACTACTGCT 420

Qy       2194 AAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCT 2253
Db        421 AAGGAGAATTATGACAATGCTAAATTGTCCTGTAGGAACCACAATGCCTTTTTGGCTTCC 480

Qy       2254 CTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCT 2313
Db        481 CTCACATCCCAGAAGAAGGTGGAGTTTGTCCTTAAGCAGCTTCGATTAATGCAATCATCT 540

Qy       2314 CAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCC 2373
Db        541 CAAAGTATGTCCAAGCTCACTCTGACTCCATGGGTTGGTCTTCGGAAGATCAATGTGTCT 600

Qy       2374 TACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCT 2433
Db        601 TACTGGTGCTGGGAGGATATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCT 660

Qy       2434 GAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTC-GGGGACTGAA 2492
Db        661 GAGCCCAGTGATGCTGGCTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGGATTAAA 720

Qy       2493 GGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAG 2552
Db        721 GGCTGCAACCTGCATCAACCCTCTCAATGGCAGCGTCTGTGAAAAGGCTGCAAACCACAG 780

Qy       2553 TGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAG-CATGTGGAGATTGCACCAGCG 2611
Db        781 TGCCAAGCAGTGCCGGACACCATGTGCCCTGCGGACANCCTTGTGGCCAGTGCACTAGCA 840

Qy       2612 GCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATG 2671
Db        841 CCAGCTCGGAGTGCATGTGGTGCAGTAACATGAAGCAATGTGGGGACTCCAATGCCTACG 900

Qy       2672 TGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTG 2731
Db        901 TGGCCTCCTTCCNTTTTGNCCAANGNATTGAAAGGTATACGATGAGCAGCTGTCCAACTG 960

Qy       2732 AAAA 2735
Db        961 AAAA 964

RESULT 6
BG678679
LOCUS       BG678679                933 bp    mRNA    linear   EST 01-MAY-2001
DEFINITION  602624477F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4749427 5',
            mRNA sequence.
ACCESSION   BG678679
VERSION     BG678679.1  GI:13910076
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 933)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: James Cleaver, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10602  row: j  column: 20
            High quality sequence stop: 762.
FEATURES             Location/Qualifiers
     source          1. .933
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4749427"
                     /tissue_type="squamous cell carcinoma"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NCI_CGAP_Skn4"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.5kb.  Library  constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
```

ORIGIN

Query Match 18.7%; Score 758.8; DB 12; Length 933;
Best Local Similarity 98.4%; Pred. No. 1e-181;
Matches 777; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

```
Qy       2775 ACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGG 2834
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 ACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGG 61

Qy       2835 TTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACA 2894
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 TTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACA 121

Qy       2895 GCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCA 2954
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCA 181

Qy       2955 CTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAA 3014
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 CTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAA 241

Qy       3015 GTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGA 3074
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 GTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGA 301

Qy       3075 TCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAA 3134
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAA 361

Qy       3135 CACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCT 3194
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 CACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCT 421

Qy       3195 ATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCT 3254
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 ATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCT 481

Qy       3255 TCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGC 3314
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 TCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGC 541

Qy       3315 TATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGC 3374
             ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        542 TATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTCGGACATGTTCATCAATGC 601

Qy       3375 CTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGC 3434
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 CTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGC 661

Qy       3435 TGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTC 3494
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 TGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTC 721

Qy       3495 TAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAA 3554
             |||||||||||  |||| |||||| || |||||||||||||| ||| ||||||||||||
Db        722 TAATGAGAAGTCCGATTCTCGCAAGCA-CCAAATATCACTTTCTGTGTGTATGTCAGTAA 780

Qy       3555 TTTCACCTGG 3564
             ||   || |
Db        781 TTGTCACTTG 790
```

RESULT 7
CB520374
LOCUS CB520374 840 bp mRNA linear EST 09-JUL-2003
DEFINITION UI-M-GI0-cei-1-18-0-UI.r1 NIH_BMAP_GI0 Mus musculus cDNA clone IMAGE:6840067 5', mRNA sequence.
ACCESSION CB520374
VERSION CB520374.1 GI:29353729
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 840)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
Seq primer: pYX-5.
FEATURES Location/Qualifiers
source 1. .840
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6840067"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_GI0"
/note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel.First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site .Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector . The library tag sequence located between the Not I site and the polyA tail is AGCGAGACAG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."
ORIGIN

Query Match 17.2%; Score 699.6; DB 14; Length 840;
Best Local Similarity 90.1%; Pred. No. 1.2e-166;
Matches 758; Conservative 0; Mismatches 82; Indels 1; Gaps 1;

```
Qy       1929 CAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGG 1988
             |||||||||||||||||||||| || ||||||||   |||||||||||||| |||||||
Db          1 CAATGACTGCCACTGGTGCAATGATCACTGTGTCCCTGTGAACCACAGCTGCACAGAAGG 60

Qy       1989 CCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAA 2048
             |||||||||||||    | |||||||| ||||||||||||||||||| |||||||||| ||
Db         61 CCAGATCTCCATTGCCAAGTATGAGAGTTGCCCCAAGGATAACCCCATGTACTACTGCAA 120

Qy       2049 CAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAA 2108
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TAAGAAAACCAGCTGCAGGAGCTGTGCCCTAGACCAGAACTGCCAGTGGGAGCCCCGGAA 180

Qy       2109 TCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAA 2168
             ||| |||||||| |||||||| |||||||||||||| |||||||||||||||||||||||
Db        181 TCAAGAGTGCATCGCCCTGCCGGAAAATATCTGTGGCAATGGCTGGCATTTGGTTGGAAA 240

Qy       2169 CTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAG 2228
             ||| ||| |||||||| |||||||| |||||||||||||||||||||||||||| ||||||
Db        241 CTCGTGTCTGAAAATCACTACTGCTAAGGAGAATTATGACAATGCTAAATTGTCCTGTAG 300

Qy       2229 GAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAA 2288
             |||||||||||| |||||||||| || ||| |||||||||||||| || ||||||||||
Db        301 GAACCACAATGCCTTTTTGGCTTCCCTCACATCCCAGAAGAAGGTGGAGTTTGTCCTTAA 360
```

```
Qy       2289 GCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGT 2348
             |||||| ||| ||||||| |||||||| || ||||||||||||||   | || ||||||||
Db        361 GCAGCTTCGATTAATGCAATCATCTCAAAGTATGTCCAAGCTCACTCTGACTCCATGGGT 420

Qy       2349 CGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAA 2408
              || |||||||||||||||||||| |||||||||||||| |||||||| ||||| |||||
Db        421 TGGTCTTCGGAAGATCAATGTGTCTTACTGGTGCTGGGAGGATATGTCTCCATTCACAAA 480

Qy       2409 TAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATC 2468
             |||||| || |||||||||| |||||||||||||||||| |||||||| || || ||
Db        481 TAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGTGATGCTGGCTTCTGTGGGATCTTGTC 540

Qy       2469 AGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGT 2528
             ||| || |||||||||||| | ||||||||||||||||||||||| |||||||| || ||
Db        541 AGAGCCTAGTACTCGGGGATTAAAGGCTGCAACCTGCATCAACCCTCTCAATGGCAGCGT 600

Qy       2529 CTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGAC 2588
             |||||||||||||||| |||||||||| |||||||||||||||||||||||||| || ||||
Db        601 CTGTGAAAGGCCTGCANACCACAGTGCCAAGCAGTGCCGGACACCATGTGCCCTGCGGAC 660

Qy       2589 AGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCA 2648
             ||| ||||| || ||||| ||| ||||||| |||||||||||||||     |||||||||
Db        661 AGCGTGTGGCGAGTGCACTAGCAGCAGCTCGGAGTGCATGTGGTGCNAGTACATGAAGCA 720

Qy       2649 GTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTA 2708
             |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db        721 GTGTGTGGACTCCAATGCCTACGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTA 780

Qy       2709 TACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTT 2768
             |||||||||| |||||| |||| | |||| || |||||||||||||| || ||||||||
Db        781 TACGATGAGCAGCTGCCCACCTG-ANATTGCTCTGGCTACTGTACCTGCAGCCATTGCTT 839

Qy       2769 G 2769
              |
Db        840 G 840
```

RESULT 8
BM783739
LOCUS BM783739 681 bp mRNA linear EST 05-MAR-2002
DEFINITION K-EST0061730 S5SNU484 Homo sapiens cDNA clone S5SNU484-34-H05 5', mRNA sequence.
ACCESSION BM783739
VERSION BM783739.1 GI:19131971
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 681)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 34 row: H column: 05
High quality sequence stop: 681.
FEATURES Location/Qualifiers
source 1. .681
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S5SNU484-34-H05"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/cell_line="SNU-484"
/lab_host="Top10F'"
/clone_lib="S5SNU484"
/note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI; Site_2: NotI; The poly (A)+ RNA was decapped with tabacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method."
ORIGIN

Query Match 16.7%; Score 681; DB 12; Length 681;
Best Local Similarity 100.0%; Pred. No. 6.1e-162;
Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2290 CAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTC 2349
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTC 60

Qy       2350 GGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAAT 2409
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAAT 120

Qy       2410 AGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCA 2469
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCA 180

Qy       2470 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTC 2529
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTC 240

Qy       2530 TGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACA 2589
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACA 300

Qy       2590 GCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 2649
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 360

Qy       2650 TGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 2709
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 420

Qy       2710 ACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTG 2769
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTG 480

Qy       2770 GAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATA 2829
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATA 540

Qy       2830 GAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTAT 2889
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTAT 600

Qy       2890 CCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 2949
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 660

Qy       2950 ATTCACTGTCCAGCTTGCCAA 2970
             |||||||||||||||||||||
Db        661 ATTCACTGTCCAGCTTGCCAA 681
```

RESULT 9
BM783788

```
LOCUS       BM783788                683 bp    mRNA    linear   EST 05-MAR-2002
DEFINITION  K-EST0061794 S5SNU484 Homo sapiens cDNA clone S5SNU484-36-E09 5',
            mRNA sequence.
ACCESSION   BM783788
VERSION     BM783788.1  GI:19132020
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 683)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 36  row: E  column: 09
            High quality sequence stop: 683.
FEATURES             Location/Qualifiers
     source          1. .683
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="S5SNU484-36-E09"
                     /sex="M"
                     /tissue_type="Stomach"
                     /cell_type="Epithelial"
                     /cell_line="SNU-484"
                     /lab_host="Top10F'"
                     /clone_lib="S5SNU484"
                     /note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI;
                     Site_2: NotI; The poly (A)+ RNA was decapped with tabacco
                     acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
                     including EcoRI site by treatment of T4 RNA ligase. The
                     first strand cDNA was synthesized from oligo dT-selected
                     mRNA by priming with dT-tailed vector. The dT-tailed
                     vector was adjusted to have about 60nt. The cDNA vector
                     was circularized with E. coli DNA ligase after digestion
                     of EcoRI which site is also included in vector. An RNA
                     strand converted to a DNA strand by Okayama-Berg method.
                     The obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation
                     method."
ORIGIN

  Query Match             16.7%;  Score 680.4;  DB 12;  Length 683;
  Best Local Similarity   99.7%;  Pred. No. 8.7e-162;
  Matches  681;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2290 CAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTC 2349
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTC 60

Qy       2350 GGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAAT 2409
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAAT 120

Qy       2410 AGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCA 2469
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCA 180

Qy       2470 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTC 2529
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTC 240

Qy       2530 TGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACA 2589
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACA 300

Qy       2590 GCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 2649
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 360

Qy       2650 TGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 2709
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 420

Qy       2710 ACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTG 2769
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTG 480

Qy       2770 GAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATA 2829
             ||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||
Db        481 GAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGNCATACTGGCAAAGGGAAATGCATA 540

Qy       2830 GAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTAT 2889
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTAT 600

Qy       2890 CCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 2949
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 660

Qy       2950 ATTCACTGTCCAGCTTGCCAATG 2972
             |||||||||||||||||||||||
Db        661 ATTCACTGTCCAGCTTGCCAATG 683


RESULT 10
BM783258
LOCUS       BM783258                680 bp    mRNA    linear   EST 05-MAR-2002
DEFINITION  K-EST0061157 S5SNU484 Homo sapiens cDNA clone S5SNU484-31-H08 5',
            mRNA sequence.
ACCESSION   BM783258
VERSION     BM783258.1  GI:19131490
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 680)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 31  row: H  column: 08
            High quality sequence stop: 680.
FEATURES             Location/Qualifiers
     source          1. .680
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="S5SNU484-31-H08"
                     /sex="M"
                     /tissue_type="Stomach"
                     /cell_type="Epithelial"
                     /cell_line="SNU-484"
                     /lab_host="Top10F'"
                     /clone_lib="S5SNU484"
                     /note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI;
                     Site_2: NotI; The poly (A)+ RNA was decapped with tabacco
```

acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method."

ORIGIN

Query Match 16.7%; Score 680; DB 12; Length 680;
Best Local Similarity 100.0%; Pred. No. 1.1e-161;
Matches 680; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2290 CAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTC 2349
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTC 60

Qy       2350 GGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAAT 2409
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAAT 120

Qy       2410 AGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCA 2469
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCA 180

Qy       2470 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTC 2529
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTC 240

Qy       2530 TGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACA 2589
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACA 300

Qy       2590 GCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 2649
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 360

Qy       2650 TGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 2709
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 420

Qy       2710 ACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTG 2769
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTG 480

Qy       2770 GAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATA 2829
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATA 540

Qy       2830 GAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTAT 2889
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTAT 600

Qy       2890 CCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 2949
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 660

Qy       2950 ATTCACTGTCCAGCTTGCCA 2969
             ||||||||||||||||||||
Db        661 ATTCACTGTCCAGCTTGCCA 680
```

RESULT 11
BM783773
LOCUS BM783773 680 bp mRNA linear EST 05-MAR-2002
DEFINITION K-EST0061775 S5SNU484 Homo sapiens cDNA clone S5SNU484-36-D02 5', mRNA sequence.
ACCESSION BM783773
VERSION BM783773.1 GI:19132005
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 680)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 36 row: D column: 02
High quality sequence stop: 680.
FEATURES Location/Qualifiers
source 1. .680
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S5SNU484-36-D02"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/cell_line="SNU-484"
/lab_host="Top10F'"
/clone_lib="S5SNU484"
/note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI; Site_2: NotI; The poly (A)+ RNA was decapped with tabacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method."

ORIGIN

Query Match 16.7%; Score 680; DB 12; Length 680;
Best Local Similarity 100.0%; Pred. No. 1.1e-161;
Matches 680; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2290 CAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTC 2349
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTC 60

Qy       2350 GGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAAT 2409
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAAT 120

Qy       2410 AGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCA 2469
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCA 180

Qy       2470 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTC 2529
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTC 240

Qy       2530 TGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACA 2589
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACA 300

Qy       2590 GCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 2649
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        301 GCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 360

Qy       2650 TGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 2709
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 420

Qy       2710 ACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTG 2769
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTG 480

Qy       2770 GAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATA 2829
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATA 540

Qy       2830 GAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTAT 2889
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTAT 600

Qy       2890 CCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 2949
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 660

Qy       2950 ATTCACTGTCCAGCTTGCCA 2969
             ||||||||||||||||||||
Db        661 ATTCACTGTCCAGCTTGCCA 680
```

RESULT 12
BM783718
LOCUS BM783718 677 bp mRNA linear EST 05-MAR-2002
DEFINITION K-EST0061702 S5SNU484 Homo sapiens cDNA clone S5SNU484-34-F01 5', mRNA sequence.
ACCESSION BM783718
VERSION BM783718.1 GI:19131950
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 677)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 34 row: F column: 01
High quality sequence stop: 677.
FEATURES Location/Qualifiers
source 1. .677
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S5SNU484-34-F01"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/cell_line="SNU-484"
/lab_host="Top10F'"
/clone_lib="S5SNU484"
/note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI; Site_2: NotI; The poly (A)+ RNA was decapped with tabacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method."

ORIGIN

```
  Query Match             16.6%;  Score 677;  DB 12;  Length 677;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-161;
  Matches  677;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2290 CAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTC 2349
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTC 60

Qy       2350 GGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAAT 2409
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAAT 120

Qy       2410 AGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCA 2469
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCA 180

Qy       2470 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTC 2529
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTC 240

Qy       2530 TGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACA 2589
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACA 300

Qy       2590 GCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 2649
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 360

Qy       2650 TGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 2709
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 420

Qy       2710 ACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTG 2769
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTG 480

Qy       2770 GAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATA 2829
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATA 540

Qy       2830 GAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTAT 2889
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTAT 600

Qy       2890 CCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 2949
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 660

Qy       2950 ATTCACTGTCCAGCTTG 2966
             |||||||||||||||||
Db        661 ATTCACTGTCCAGCTTG 677
```

RESULT 13
BM783821
LOCUS BM783821 677 bp mRNA linear EST 05-MAR-2002
DEFINITION K-EST0061830 S5SNU484 Homo sapiens cDNA clone S5SNU484-36-H09 5', mRNA sequence.
ACCESSION BM783821
VERSION BM783821.1 GI:19132053
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```
REFERENCE   1  (bases 1 to 677)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 36  row: H  column: 09
            High quality sequence stop: 677.
FEATURES             Location/Qualifiers
     source          1. .677
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="S5SNU484-36-H09"
                     /sex="M"
                     /tissue_type="Stomach"
                     /cell_type="Epithelial"
                     /cell_line="SNU-484"
                     /lab_host="Top10F'"
                     /clone_lib="S5SNU484"
                     /note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI;
                     Site_2: NotI; The poly (A)+ RNA was decapped with tabacco
                     acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
                     including EcoRI site by treatment of T4 RNA ligase. The
                     first strand cDNA was synthesized from oligo dT-selected
                     mRNA by priming with dT-tailed vector. The dT-tailed
                     vector was adjusted to have about 60nt. The cDNA vector
                     was circularized with E. coli DNA ligase after digestion
                     of EcoRI which site is also included in vector. An RNA
                     strand converted to a DNA strand by Okayama-Berg method.
                     The obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation
                     method."
ORIGIN

  Query Match           16.6%;  Score 674.4;  DB 12;  Length 677;
  Best Local Similarity  99.7%;  Pred. No. 2.9e-160;
  Matches  675;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2290 CAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTC 2349
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTC 60

Qy       2350 GGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAAT 2409
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAAT 120

Qy       2410 AGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCA 2469
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCA 180

Qy       2470 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTC 2529
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTC 240

Qy       2530 TGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACA 2589
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACA 300

Qy       2590 GCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 2649
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 360

Qy       2650 TGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 2709
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 420

Qy       2710 ACGATGAGCACCTGCCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTG 2769
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACGATGAGCACCTGCCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTG 480

Qy       2770 GAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATA 2829
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATA 540

Qy       2830 GAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTAT 2889
             ||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||
Db        541 GAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGNCAGCCCCTACAGGAAATTTCTAT 600

Qy       2890 CCACAGCCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 2949
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCACAGCCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 660

Qy       2950 ATTCACTGTCCAGCTTG 2966
             ||||||||||||||||
Db        661 ATTCACTGTCCAGCTTG 677


RESULT 14
BM783685
LOCUS       BM783685                673 bp    mRNA    linear   EST 05-MAR-2002
DEFINITION  K-EST0061659 S5SNU484 Homo sapiens cDNA clone S5SNU484-34-B06 5',
            mRNA sequence.
ACCESSION   BM783685
VERSION     BM783685.1  GI:19131917
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 673)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 34  row: B  column: 06
            High quality sequence stop: 673.
FEATURES             Location/Qualifiers
     source          1. .673
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="S5SNU484-34-B06"
                     /sex="M"
                     /tissue_type="Stomach"
                     /cell_type="Epithelial"
                     /cell_line="SNU-484"
                     /lab_host="Top10F'"
                     /clone_lib="S5SNU484"
                     /note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI;
                     Site_2: NotI; The poly (A)+ RNA was decapped with tabacco
                     acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
                     including EcoRI site by treatment of T4 RNA ligase. The
                     first strand cDNA was synthesized from oligo dT-selected
                     mRNA by priming with dT-tailed vector. The dT-tailed
                     vector was adjusted to have about 60nt. The cDNA vector
                     was circularized with E. coli DNA ligase after digestion
                     of EcoRI which site is also included in vector. An RNA
                     strand converted to a DNA strand by Okayama-Berg method.
                     The obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation
```

method."
ORIGIN

Query Match 16.5%; Score 673; DB 12; Length 673;
Best Local Similarity 100.0%; Pred. No. 6.7e-160;
Matches 673; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2290 CAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTC 2349
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTC 60

Qy       2350 GGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAAT 2409
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAAT 120

Qy       2410 AGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCA 2469
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCA 180

Qy       2470 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTC 2529
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTC 240

Qy       2530 TGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACA 2589
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACA 300

Qy       2590 GCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 2649
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 360

Qy       2650 TGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 2709
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 420

Qy       2710 ACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTG 2769
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTG 480

Qy       2770 GAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATA 2829
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATA 540

Qy       2830 GAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTAT 2889
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTAT 600

Qy       2890 CCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 2949
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 660

Qy       2950 ATTCACTGTCCAG 2962
             |||||||||||||
Db        661 ATTCACTGTCCAG 673
```

RESULT 15
CD644584
LOCUS CD644584 792 bp mRNA linear EST 17-JUN-2003
DEFINITION AGENCOURT_14539338 NIA Human H1 Embryonic Stem Cell cDNA Library (Long) Homo sapiens cDNA clone IMAGE:30421895 5', mRNA sequence.
ACCESSION CD644584
VERSION CD644584.1 GI:31816608
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 792)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Irene Ginis and Mahendra Rao, NIA
cDNA Library Preparation: Yulan Piao and Minoru Ko
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC c lone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM500 row: g column: 24
High quality sequence stop: 681.
FEATURES Location/Qualifiers
source 1. .792
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30421895"
/tissue_type="Embryonic Stem cells"
/cell_line="WA01"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIA Human H1 Embryonic Stem Cell cDNA Library (Long)"
/note="Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI; This is a long-transcript enriched cDNA library (Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]) from WA01 cell line . Undifferentiated human ES cell line WA01/H1 was obtained from WiCell Research Institute, Inc., Madison, Wi, cultured according to their instructions, on MEF feeders. They formed round colonies with defined edges and were positive for alkaline phosphatase, SSEA-4, OCT3, OCT4, REX1, UTF, TERT, SOX2, CX43 and CX45. They are negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLT3, SSEA-1, TUBB3, NES, GFAP, and EOMES. When confluent (18-10 days after plating), the ES cells from 4 X 6cm dishes were treated with 1 mg/ml collagenase, type IV (Invitrogen/GIBCO) for 5-10 min and gently scraped off with 5 ml pipette. RNA was purified with TRIzol Reagent from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558 (2001). [PMID:11544199]) Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT-3'] from 3.4g of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-Sal4, purified by phenol/chloroform extraction, and separated from free linkers by Centricon-100 column. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S for 25 cycles. The products were purified by phenol/chloroform extraction and Centricon-100 column. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The average insert size is about 3.6kb."
ORIGIN

Query Match 16.4%; Score 665.4; DB 14; Length 792;
Best Local Similarity 95.0%; Pred. No. 6.3e-158;
Matches 687; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

```
Qy        257 GGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAG 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         12 GGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAG 71

Qy        317 AGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTG 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 AGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTG 131

Qy        377 CTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTG 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 CTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTG 191
```

```
Qy        437 TTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 TTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTG 251

Qy        497 GATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGT 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 GATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGT 311

Qy        557 GTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTG 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 GTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTG 371

Qy        617 AAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCA 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 AAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCA 431

Qy        677 ATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAG 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 ATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAG 491

Qy        737 TTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCC 796
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        492 TTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCC 551

Qy        797 CCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATA 856
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        552 CCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATA 611

Qy        857 TGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGC 916
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        612 TGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGC 671

Qy        917 TTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTAT 976
              |||   ||    | ||  |   || || | ||| |   |   | |   ||  ||
Db        672 TTCACTAAACGTTCTGTGAACATGTGGTTGTAGATAGGTCATCTTTGGCATATACAGGAT 731

Qy        977 ACA 979
              | |
Db        732 AAA 734
```

Search completed: March 5, 2004, 12:07:55
Job time : 6449.51 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2004, 16:06:34 ; Search time 958.587 Seconds
(without alignments)
18028.269 Million cell updates/sec

Title: US-09-787-097-11
Perfect score: 4068
Sequence: 1 atggtggccgcagcggcggc.....agcctgggacctgcatctga 4068

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	4068	100.0	4068	3	AAZ52275	Aaz52275 Human mem
2	3988.4	98.0	8590	3	AAZ91920	Aaz91920 Human mah
3	3981	97.9	4290	3	AAZ52276	Aaz52276 Human mem
4	3981	97.9	8561	5	AAS72659	Aas72659 DNA encod
5	3556.4	87.4	3597	3	AAZ52274	Aaz52274 Human sol
6	3556.4	87.4	3597	5	AAS72657	Aas72657 DNA encod
7	3556.4	87.4	3597	9	ADD70997	Add70997 Human att
8	3499.8	86.0	4072	3	AAZ91921	Aaz91921 Human mah
9	3492.4	85.9	3819	3	AAZ52277	Aaz52277 Human sol
10	3353.2	82.4	8739	9	ADB53249	Adb53249 Primary r
11	3324.8	81.7	8827	3	AAZ91916	Aaz91916 Wild type
12	2810.4	69.1	3490	5	AAS72660	Aas72660 DNA encod
13	2051.8	50.4	2625	3	AAZ91922	Aaz91922 Human mah
14	1830.2	45.0	6373	3	AAZ91919	Aaz91919 Human mah
15	1481	36.4	4134	7	ABZ58710	Abz58710 Murine ho
16	1465	36.0	4140	7	ABZ58709	Abz58709 Human hom
17	1451.8	35.7	2419	3	AAZ91917	Aaz91917 Murine ma
18	1131	27.8	3609	6	ABS63434	Abs63434 Human cDN
19	873.2	21.5	3165	5	AAS78781	Aas78781 DNA encod
20	873.2	21.5	3313	5	AAS76294	Aas76294 DNA encod
21	817.8	20.1	1012	2	AAX84568	Aax84568 Biorhythm
22	812.8	20.0	6517	4	AAH98748	Aah98748 Human EST
23	773.2	19.0	6532	5	AAS86766	Aas86766 DNA encod
24	463	11.4	1051	3	AAZ91918	Aaz91918 Murine ma
25	418.4	10.3	1654	7	ABX34730	Abx34730 Human mdd
26	413.6	10.2	1383	6	AAL54064	Aal54064 cDNA of h
27	404.2	9.9	671	5	AAS72658	Aas72658 DNA encod
28	314.4	7.7	668	4	AAS02409	Aas02409 Human sec
29	278	6.8	292	3	AAA45505	Aaa45505 Human sec
c 30	277.4	6.8	481	4	AAL14686	Aal14686 Human bre
c 31	266.4	6.5	482	4	AAL23548	Aal23548 Human bre
32	253.8	6.2	982	6	ABL39784	Abl39784 Human NS
33	250.8	6.2	531	4	AAI43425	Aai43425 Probe #12
34	250.8	6.2	531	6	ABS11523	Abs11523 Human gen
c 35	246	6.0	246	6	ABK38928	Abk38928 cDNA enco
c 36	246	6.0	246	7	ACA11257	Aca11257 Human lun
c 37	246	6.0	246	7	ACA02443	Aca02443 Lung canc
38	236	5.8	657	7	ABZ58714	Abz58714 Upregulat
39	234.4	5.8	207433	5	ABZ72040	Abz72040 Gene 216
40	234.4	5.8	207433	7	ABX74891	Abx74891 BAC1098L2
41	202.8	5.0	90050	3	AAZ91925	Aaz91925 Wild type
42	201	4.9	4164	4	ABL09747	Abl09747 Drosophil
43	193	4.7	193	4	AAI56405	Aai56405 Probe #25
44	193	4.7	193	6	ABS23881	Abs23881 Human gen
c 45	191	4.7	10315	4	ABL09746	Abl09746 Drosophil

ALIGNMENTS

RESULT 1
AAZ52275
ID AAZ52275 standard; cDNA; 4068 BP.
XX
AC AAZ52275;
XX
DT 18-JUL-2000 (first entry)
XX
DE Human membrane attractin-1 cDNA.
XX
KW Human; membrane attractin-1; immune response; macrophage; monocyte;
KW T cell; immunostimulant; immunosuppressed patient; cancer;
KW immunodeficiency syndrome; transplant; autoimmune disease; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1. .4068
FT /*tag= a
FT /product= "Membrane attractin-1"
XX
PN WO200015651-A1.
XX
PD 23-MAR-2000.
XX
PF 14-SEP-1999; 99WO-US020948.
XX
PR 14-SEP-1998; 98US-0100137P.
XX
PA (DAND) DANA FARBER CANCER INST INC.
XX
PI Duke-Cohan JS, Schlossman SF;
XX
DR WPI; 2000-271373/23.
DR P-PSDB; AAY70690.
XX
PT Isolated nucleic acids encoding human attractin polypeptides useful for
PT enhancing immune responses.
XX
PS Claim 3; Fig 10; 120pp; English.
XX
CC The patent discloses four forms of human attractin polypeptides which
CC enhance immune response by promoting macrophage and monocyte spreading in
CC the presence of T cells. These include soluble attractin-1 and -2 and
CC membrane attractin-1 and -2. These various forms of attractin are encoded
CC by alternatively spliced mRNA molecule transcribed from a single gene.

CC The present sequence is a cDNA encoding human membrane attractin-1.
CC Membrane attractin differs from soluble attractin in having a
CC transmembrane domain and a cytoplasmic domain at the C-terminal. This
CC sequence can be used to enhance immune response in immunosuppressed
CC patients such as those undergoing chemo- and radio-therapy treatment for
CC cancer or those suffering from common variable immunodeficiency syndrome.
CC The proteins may also be used to screen modulators (agonists and
CC antagonists) of immune response which may also be used to regulate immune
CC reactions. Attractin antibodies can be used to inhibit immune response in
CC transplant recipients or patients afflicted with autoimmune disease
XX
SQ Sequence 4068 BP; 1042 A; 948 C; 1027 G; 1051 T; 0 U; 0 Other;

Query Match 100.0%; Score 4068; DB 3; Length 4068;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4068; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 60

Qy         61 GCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGC 120

Qy        121 ACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGC 180

Qy        181 CGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATAC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATAC 240

Qy        241 AAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTC 300

Qy        301 AATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATT 360

Qy        361 TATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAG 420

Qy        421 ACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCT 480

Qy        481 GCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGC 540

Qy        541 TCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCT 600

Qy        601 GAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCT 660

Qy        661 CATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAG 720

Qy        721 GGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATAT 780

Qy        781 TCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGG 840

Qy        841 GTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTT 900

Qy        901 GCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGT 960

Qy        961 CATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACT 1020

Qy       1021 GGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTG 1080

Qy       1081 ACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTG 1140

Qy       1141 AAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATA 1200

Qy       1201 AGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGT 1260

Qy       1261 GCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTA 1320

Qy       1321 TACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTC 1380

Qy       1381 TACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGT 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGT 1440

Qy       1441 TACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACAC 1500

Qy       1501 AATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGAC 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 AATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGAC 1560

Qy       1561 ATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGA 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 ATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGA 1620

Qy       1621 TTTGGCCATTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGT 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TTTGGCCATTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGT 1680

Qy       1681 CTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAA 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAA 1740

Qy       1741 GCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAG 1800

Qy       1801 TGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTT 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 TGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTT 1860

Qy       1861 TCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACA 1920

Qy       1921 GCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGC 1980
```

```
Db       1921 GCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGC 1980
Qy       1981 TCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTAC 2040
Db       1981 TCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTAC 2040
Qy       2041 TACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAG 2100
Db       2041 TACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAG 2100
Qy       2101 CCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTG 2160
Db       2101 CCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTG 2160
Qy       2161 GTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTG 2220
Db       2161 GTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTG 2220
Qy       2221 TTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTT 2280
Db       2221 TTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTT 2280
Qy       2281 GTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACC 2340
Db       2281 GTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACC 2340
Qy       2341 CCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCA 2400
Db       2341 CCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCA 2400
Qy       2401 TTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGA 2460
Db       2401 TTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGA 2460
Qy       2461 ATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAAT 2520
Db       2461 ATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAAT 2520
Qy       2521 GGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCC 2580
Db       2521 GGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCC 2580
Qy       2581 TTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAAC 2640
Db       2581 TTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAAC 2640
Qy       2641 ATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATG 2700
Db       2641 ATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATG 2700
Qy       2701 GAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGT 2760
Db       2701 GAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGT 2760
Qy       2761 CATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGG 2820
Db       2761 CATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGG 2820
Qy       2821 AAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGA 2880
Db       2821 AAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGA 2880
Qy       2881 AATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAAC 2940
Db       2881 AATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAAC 2940
Qy       2941 TGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAG 3000
Db       2941 TGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAG 3000
Qy       3001 AGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCT 3060
```

```
Db       3001 AGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCT 3060
Qy       3061 GGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCAC 3120
Db       3061 GGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCAC 3120
Qy       3121 GCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGG 3180
Db       3121 GCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGG 3180
Qy       3181 GACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACA 3240
Db       3181 GACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACA 3240
Qy       3241 TGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGAT 3300
Db       3241 TGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGAT 3300
Qy       3301 CGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGAC 3360
Db       3301 CGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGAC 3360
Qy       3361 ATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCA 3420
Db       3361 ATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCA 3420
Qy       3421 GCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTAC 3480
Db       3421 GCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTAC 3480
Qy       3481 AAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTT 3540
Db       3481 AAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTT 3540
Qy       3541 GTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGATTGCCTTCTCTCAGCACAGC 3600
Db       3541 GTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGATTGCCTTCTCTCAGCACAGC 3600
Qy       3601 AATTTTATGGACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCTTTGCTC 3660
Db       3601 AATTTTATGGACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCTTTGCTC 3660
Qy       3661 CTGGTGGCTGCTGTGGTTTGGAAGATCAAACAAAGTTGTTGGGCCTCCAGACGTAGAGAG 3720
Db       3661 CTGGTGGCTGCTGTGGTTTGGAAGATCAAACAAAGTTGTTGGGCCTCCAGACGTAGAGAG 3720
Qy       3721 CAACTTCTTCGAGAGATGCAACAGATGGCCAGCCGTCCCTTTGCCTCTGTAAATGTCGCC 3780
Db       3721 CAACTTCTTCGAGAGATGCAACAGATGGCCAGCCGTCCCTTTGCCTCTGTAAATGTCGCC 3780
Qy       3781 TTGGAAACAGATGAGGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAA 3840
Db       3781 TTGGAAACAGATGAGGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAA 3840
Qy       3841 CCCATTGCACTGGAGCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGG 3900
Db       3841 CCCATTGCACTGGAGCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGG 3900
Qy       3901 CTCCCTCGAGGCCTGGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGC 3960
Db       3901 CTCCCTCGAGGCCTGGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGC 3960
Qy       3961 GCCCTGGTGGACATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTG 4020
Db       3961 GCCCTGGTGGACATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTG 4020
Qy       4021 AGAAACCGGAAGCAGCAGCCCCCTGCACAGCCTGGGACCTGCATCTGA 4068
Db       4021 AGAAACCGGAAGCAGCAGCCCCCTGCACAGCCTGGGACCTGCATCTGA 4068

RESULT 2
AAZ91920
```

ID AAZ91920 standard; cDNA; 8590 BP.
XX
AC AAZ91920;
XX
DT 08-JUN-2000 (first entry)
XX
DE Human mahogany protein coding sequence #2.
XX
KW Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;
KW weight regulation; cell therapy; body weight disorder; cachexia;
KW anorexia; hyperpigmentation; increased metabolic rate disorder;
KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.
XX
OS Homo sapiens.
XX
PN WO200005373-A2.
XX
PD 03-FEB-2000.
XX
PF 21-JUL-1999; 99WO-US016484.
XX
PR 21-JUL-1998; 98US-0093630P.
PR 20-OCT-1998; 98US-0104978P.
PR 05-FEB-1999; 99US-00245041.
XX
PA (MILL-) MILLENIUM PHARM INC.
XX
PI Moore K, Nagle DL;
XX
DR WPI; 2000-195103/17.
DR P-PSDB; AAY81807.
XX
PT New human and murine mahogany genes, useful, e.g. for diagnosis and
PT treatment of body weight disorders.
XX
PS Claim 1; Fig 18a; 188pp; English.
XX
CC This sequence represents a human mahogany gene of the invention. The
CC mahogany genes are used: (i) to produce recombinant mahogany (mg)
CC proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming
CC therapeutics; (iii) as a source of diagnostic probes and primers for
CC detecting expression of mg genes or mutations, regulatory defects, in
CC this gene, or for isolation of related sequences; and (iv) in (cell-
CC based) gene therapy. (II) are used to raise specific antibodies (Ab); to
CC identify other (extra)cellular products involved in weight regulation,
CC and to screen for agents that disrupt interaction between (II) and other
CC macromolecules. The Ab are used to detect abnormal levels (or function)
CC of (II) (for diagnosis, prognosis or monitoring of treatment); to
CC evaluate (II)-expressing cells intended for cell therapy, and as
CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the
CC mg polypeptide) are used to identify agents (A) that modulate mg
CC activity. (A) are potentially useful for the treatment of body weight
CC disorders, particularly obesity, cachexia or anorexia, or other
CC conditions associated with the mg gene such as hyperpigmentation,
CC hyperphagia and disorders that result in increased metabolic rate
XX
SQ Sequence 8590 BP; 2275 A; 1990 C; 2061 G; 2264 T; 0 U; 0 Other;

Query Match 98.0%; Score 3988.4; DB 3; Length 8590;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 4013; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```
Qy        15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74
Db       234 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAA 293

Qy        75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134
Db       294 GGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 353

Qy       135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194
Db       354 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 413

Qy       195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254
Db       414 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 473

Qy       255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314
Db       474 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 533

Qy       315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374
Db       534 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 593

Qy       375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434
Db       594 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 653

Qy       435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494
Db       654 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 713

Qy       495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554
Db       714 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 773

Qy       555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614
Db       774 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 833

Qy       615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 674
Db       834 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 893

Qy       675 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 734
Db       894 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 953

Qy       735 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 794
Db       954 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 1013

Qy       795 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 854
Db      1014 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 1073

Qy       855 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 914
Db      1074 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 1133

Qy       915 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 974
Db      1134 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 1193

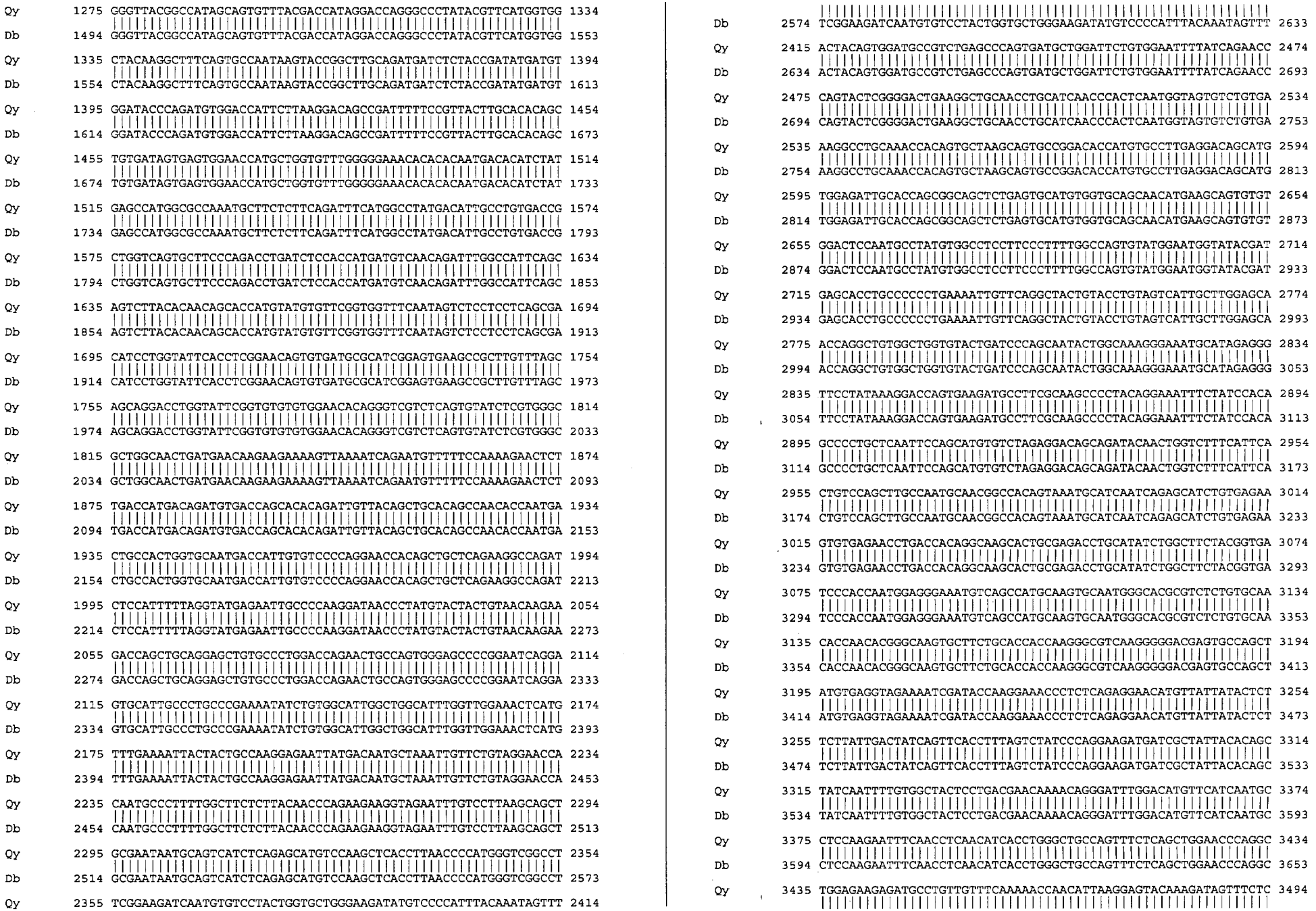
Qy       975 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1034
Db      1194 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1253

Qy      1035 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1094
Db      1254 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1313

Qy      1095 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1154
Db      1314 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1373

Qy      1155 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1214
Db      1374 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1433

Qy      1215 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1274
Db      1434 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1493
```


Qy 1275 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1334
Db 1494 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1553
Qy 1335 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1394
Db 1554 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1613
Qy 1395 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1454
Db 1614 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1673
Qy 1455 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1514
Db 1674 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1733
Qy 1515 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1574
Db 1734 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1793
Qy 1575 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGC 1634
Db 1794 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGC 1853
Qy 1635 AGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGA 1694
Db 1854 AGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGA 1913
Qy 1695 CATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGC 1754
Db 1914 CATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGC 1973
Qy 1755 AGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGC 1814
Db 1974 AGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGC 2033
Qy 1815 GCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCT 1874
Db 2034 GCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCT 2093
Qy 1875 TGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGA 1934
Db 2094 TGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGA 2153
Qy 1935 CTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGAT 1994
Db 2154 CTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGAT 2213
Qy 1995 CTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAA 2054
Db 2214 CTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAA 2273
Qy 2055 GACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGA 2114
Db 2274 GACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGA 2333
Qy 2115 GTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATG 2174
Db 2334 GTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATG 2393
Qy 2175 TTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCA 2234
Db 2394 TTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCA 2453
Qy 2235 CAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCT 2294
Db 2454 CAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCT 2513
Qy 2295 GCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCT 2354
Db 2514 GCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCT 2573
Qy 2355 TCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTT 2414
Db 2574 TCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTT 2633
Qy 2415 ACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACC 2474
Db 2634 ACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACC 2693
Qy 2475 CAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGA 2534
Db 2694 CAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGA 2753
Qy 2535 AAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATG 2594
Db 2754 AAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATG 2813
Qy 2595 TGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGT 2654
Db 2814 TGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGT 2873
Qy 2655 GGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGAT 2714
Db 2874 GGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGAT 2933
Qy 2715 GAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCA 2774
Db 2934 GAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCA 2993
Qy 2775 ACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGG 2834
Db 2994 ACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGG 3053
Qy 2835 TTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACA 2894
Db 3054 TTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACA 3113
Qy 2895 GCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCA 2954
Db 3114 GCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCA 3173
Qy 2955 CTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAA 3014
Db 3174 CTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAA 3233
Qy 3015 GTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGA 3074
Db 3234 GTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGA 3293
Qy 3075 TCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAA 3134
Db 3294 TCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAA 3353
Qy 3135 CACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCT 3194
Db 3354 CACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCT 3413
Qy 3195 ATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCT 3254
Db 3414 ATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCT 3473
Qy 3255 TCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGC 3314
Db 3474 TCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGC 3533
Qy 3315 TATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGC 3374
Db 3534 TATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGC 3593
Qy 3375 CTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGC 3434
Db 3594 CTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGC 3653
Qy 3435 TGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTC 3494

```
Db       3654 TGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTC 3713

Qy       3495 TAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAA 3554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3714 TAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAA 3773

Qy       3555 TTTCACCTGGCCCATCAAAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCT 3614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3774 TTTCACCTGGCCCATCAAAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCT 3833

Qy       3615 GGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGT 3674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3834 GGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGT 3893

Qy       3675 GGTTTGGAAGATCAAACAAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGA 3734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3894 GGTTTGGAAGATCAAACAAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGA 3953

Qy       3735 GATGCAACAGATGGCCAGCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGA 3794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3954 GATGCAACAGATGGCCAGCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGA 4013

Qy       3795 GGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGA 3854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4014 GGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGA 4073

Qy       3855 GCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCT 3914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4074 GCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCT 4133

Qy       3915 GGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACAT 3974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4134 GGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACAT 4193

Qy       3975 TTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCA 4034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4194 TTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCA 4253

Qy       4035 GCAGCCCCCTGCACAGCCTGGGACCTGCATCTGA 4068
              ||||||||||||||||||||||||||||||||||
Db       4254 GCAGCCCCCTGCACAGCCTGGGACCTGCATCTGA 4287


RESULT 3
AAZ52276
ID   AAZ52276 standard; cDNA; 4290 BP.
XX
AC   AAZ52276;
XX
DT   18-JUL-2000  (first entry)
XX
DE   Human membrane attractin-2 cDNA.
XX
KW   Human; membrane attractin-2; immune response; macrophage; monocyte;
KW   T cell; immunostimulant; immunosuppressed patient; cancer;
KW   immunodeficiency syndrome; transplant; autoimmune disease; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .4290
FT                   /*tag=  a
FT                   /product= "Membrane attractin-2"
XX
PN   WO200015651-A1.
XX
PD   23-MAR-2000.
XX
PF   14-SEP-1999;   99WO-US020948.
XX
PR   14-SEP-1998;   98US-0100137P.
XX
PA   (DAND ) DANA FARBER CANCER INST INC.
XX
PI   Duke-Cohan JS,  Schlossman SF;
XX
DR   WPI; 2000-271373/23.
DR   P-PSDB; AAY70691.
XX
PT   Isolated nucleic acids encoding human attractin polypeptides useful for
PT   enhancing immune responses.
XX
PS   Claim 3; Fig 14; 120pp; English.
XX
CC   The patent discloses four forms of human attractin polypeptides which
CC   enhance immune response by promoting macrophage and monocyte spreading in
CC   the presence of T cells. These include soluble attractin-1 and -2 and
CC   membrane attractin-1 and -2. These various forms of attractin are encoded
CC   by alternatively spliced mRNA molecule transcribed from a single gene.
CC   The present sequence is a cDNA encoding human membrane attractin-2.
CC   Membrane attractin differs from soluble attractin in having a
CC   transmembrane domain and a cytoplasmic domain at the C-terminal.
CC   Attractin-2 differs from attractin-1 in having a 74 amino acid insert in
CC   the N-terminal. This sequence can be used to enhance immune response in
CC   immunosuppressed patients such as those undergoing chemo- and radio-
CC   therapy treatment for cancer or those suffering from common variable
CC   immunodeficiency syndrome. The proteins may also be used to screen
CC   modulators (agonists and antagonists) of immune response which may also
CC   be used to regulate immune reactions. Attractin antibodies can be used to
CC   inhibit immune response in transplant recipients or patients afflicted
CC   with autoimmune disease
XX
SQ   Sequence 4290 BP; 1060 A; 1025 C; 1120 G; 1085 T; 0 U; 0 Other;

  Query Match             97.9%;  Score 3981;  DB 3;  Length 4290;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 3996;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy         48 GGCGGCGACGGCAGCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGTGTCAACGGCGGTCG 107
              | |||   | || ||   || || || |   |  |  | || ||||||||||||||||||
Db        270 GTCGGGCTCAGCCGCAGCCGAGGCCAAGGAATGTGACCGGCCCTGTGTCAACGGCGGTCG 329

Qy        108 CTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCA 167
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        330 CTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCA 389

Qy        168 GCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGG 227
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        390 GCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGG 449

Qy        228 AAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAAT 287
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        450 AAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAAT 509

Qy        288 GAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGA 347
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        510 GAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGA 569

Qy        348 TGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAG 407
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        570 TGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAG 629

Qy        408 AGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTT 467
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        630 AGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTT 689

Qy        468 TTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTG 527
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        690 TTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTG 749

Qy        528 TCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGA 587
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        750 TCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGAAACTGTTGA 809

Qy        588 ATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA 647
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        810 ATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA 869
Qy        648 CTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTT 707
Db        870 CTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTT 929
Qy        708 CTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGAC 767
Db        930 CTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGAC 989
Qy        768 TCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGG 827
Db        990 TCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGG 1049
Qy        828 AAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCT 887
Db       1050 AAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCT 1109
Qy        888 AGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGT 947
Db       1110 AGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGT 1169
Qy        948 TGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAA 1007
Db       1170 TGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAA 1229
Qy       1008 AATTGATTCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTC 1067
Db       1230 AATTGATCCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTC 1289
Qy       1068 ATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACA 1127
Db       1290 ATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACA 1349
Qy       1128 CATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCT 1187
Db       1350 CATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCT 1409
Qy       1188 CTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATT 1247
Db       1410 CTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATT 1469
Qy       1248 ACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAG 1307
Db       1470 ACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAG 1529
Qy       1308 GACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCT 1367
Db       1530 GACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCT 1589
Qy       1368 TGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAG 1427
Db       1590 TGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAG 1649
Qy       1428 CCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGG 1487
Db       1650 CCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGG 1709
Qy       1488 GGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTT 1547
Db       1710 GGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTT 1769
Qy       1548 CATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCA 1607
Db       1770 CATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCA 1829
Qy       1608 TGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGG 1667
Db       1830 TGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGG 1889
Qy       1668 TGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGC 1727
Db       1890 TGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGC 1949

Qy       1728 GCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACAC 1787
Db       1950 GCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACAC 2009
Qy       1788 AGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAA 1847
Db       2010 AGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAA 2069
Qy       1848 ATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTG 1907
Db       2070 ATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTG 2129
Qy       1908 TTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAG 1967
Db       2130 TTACAGCTGTACAGCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAG 2189
Qy       1968 GAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGA 2027
Db       2190 GAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGA 2249
Qy       2028 TAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAA 2087
Db       2250 TAACCCCATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAA 2309
Qy       2088 CTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCAT 2147
Db       2310 CTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCAT 2369
Qy       2148 TGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGA 2207
Db       2370 TGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGA 2429
Qy       2208 CAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAA 2267
Db       2430 CAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAA 2489
Qy       2268 GAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAA 2327
Db       2490 GAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAA 2549
Qy       2328 GCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGA 2387
Db       2550 GCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGA 2609
Qy       2388 AGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGC 2447
Db       2610 AGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGC 2669
Qy       2448 TGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCAT 2507
Db       2670 TGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCAT 2729
Qy       2508 CAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCG 2567
Db       2730 CAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCG 2789
Qy       2568 GACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCAT 2627
Db       2790 GACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCAT 2849
Qy       2628 GTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTT 2687
Db       2850 GTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTT 2909
Qy       2688 TGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTA 2747
Db       2910 TGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTA 2969
Qy       2748 CTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAA 2807
Db       2970 CTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAA 3029
```

```
Qy       2808 TACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCA 2867
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3030 TACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCA 3089

Qy       2868 AGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGA 2927
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3090 AGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGA 3149

Qy       2928 CAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAA 2987
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3150 CAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAA 3209

Qy       2988 ATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGA 3047
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3210 ATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGA 3269

Qy       3048 GACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAA 3107
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3270 GACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAA 3329

Qy       3108 GTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAA 3167
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3330 GTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAA 3389

Qy       3168 GGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCC 3227
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3390 GGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCC 3449

Qy       3228 TCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATC 3287
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3450 TCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATC 3509

Qy       3288 CCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAA 3347
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3510 CCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAA 3569

Qy       3348 CAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGC 3407
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3570 CAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGC 3629

Qy       3408 TGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAA 3467
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3630 TGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAA 3689

Qy       3468 CATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAA 3527
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3690 CATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAA 3749

Qy       3528 TATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGATTGCCTT 3587
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3750 TATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGATTGCCTT 3809

Qy       3588 CTCTCAGCACAGCAATTTTATGGACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTT 3647
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3810 CTCTCAGCACAGCAATTTTATGGACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTT 3869

Qy       3648 CCTCTCTTTGCTCCTGGTGGCTGCTGTGGTTTGGAAGATCAAACAAAGTTGTTGGGCCTC 3707
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3870 CCTCTCTTTGCTCCTGGTGGCTGCTGTGGTTTGGAAGATCAAACAAAGTTGTTGGGCCTC 3929

Qy       3708 CAGACGTAGAGAGCAACTTCTTCGAGAGATGCAACAGATGGCCAGCCGTCCCTTTGCCTC 3767
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3930 CAGACGTAGAGAGCAACTTCTTCGAGAGATGCAACAGATGGCCAGCCGTCCCTTTGCCTC 3989

Qy       3768 TGTAAATGTCGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTATTGGGGGGAGTATAAA 3827
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3990 TGTAAATGTCGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTATTGGGGGGAGTATAAA 4049

Qy       3828 GACTGTTCCCAAACCCATTGCACTGGAGCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTC 3887
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4050 GACTGTTCCCAAACCCATTGCACTGGAGCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTC 4109

Qy       3888 TGTGTTTGTGAGGCTCCCTCGAGGCCTGGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCT 3947
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4110 TGTGTTTGTGAGGCTCCCTCGAGGCCTGGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCT 4169

Qy       3948 TGCTGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAA 4007
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4170 TGCTGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAA 4229

Qy       4008 GTCAGGAGCCGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGCCTGGGACCTGCATCTG 4067
              - ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4230 GTCAGGAGCCGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGCCTGGGACCTGCATCTG 4289

Qy       4068 A 4068
              |
Db       4290 A 4290
```

RESULT 4
AAS72659
ID AAS72659 standard; cDNA; 8561 BP.
XX
AC AAS72659;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #8463.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG08472.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 8463; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in

```
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 8561 BP; 2246 A; 1983 C; 2065 G; 2267 T; 0 U; 0 Other;

  Query Match            97.9%;  Score 3981;  DB 5;  Length 8561;
  Best Local Similarity  99.4%;  Pred. No. 0;
  Matches 3996;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy         48 GGCGGCGACGGCAGCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGTGTCAACGGCGGTCG 107
Db        270 GTCGGGCTCAGCCGCAGCCGAGGCCAAGGAATGTGACCGGCCCTGTGTCAACGGCGGTCG 329

Qy        108 CTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCA 167
Db        330 CTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCA 389

Qy        168 GCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGG 227
Db        390 GCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGG 449

Qy        228 AAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAAT 287
Db        450 AAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAAT 509

Qy        288 GAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGA 347
Db        510 GAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGA 569

Qy        348 TGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAG 407
Db        570 TGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAG 629

Qy        408 AGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTT 467
Db        630 AGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTT 689

Qy        468 TTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTG 527
Db        690 TTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTG 749

Qy        528 TCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGA 587
Db        750 TCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGAAACTGTTGA 809

Qy        588 ATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA 647
Db        810 ATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA 869

Qy        648 CTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTT 707
Db        870 CTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTT 929

Qy        708 CTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGAC 767
Db        930 CTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGAC 989

Qy        768 TCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGG 827
Db        990 TCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGG 1049

Qy        828 AAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCT 887
Db       1050 AAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCT 1109

Qy        888 AGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGT 947
Db       1110 AGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGT 1169

Qy        948 TGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAA 1007
Db       1170 TGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAA 1229

Qy       1008 AATTGATTCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTC 1067
Db       1230 AATTGATCCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTC 1289

Qy       1068 ATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACA 1127
Db       1290 ATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACA 1349

Qy       1128 CATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCT 1187
Db       1350 CATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCT 1409

Qy       1188 CTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATT 1247
Db       1410 CTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATT 1469

Qy       1248 ACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAG 1307
Db       1470 ACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAG 1529

Qy       1308 GACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCT 1367
Db       1530 GACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCT 1589

Qy       1368 TGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAG 1427
Db       1590 TGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAG 1649

Qy       1428 CCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGG 1487
Db       1650 CCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGG 1709

Qy       1488 GGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTT 1547
Db       1710 GGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTT 1769

Qy       1548 CATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCA 1607
Db       1770 CATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCA 1829

Qy       1608 TGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGG 1667
Db       1830 TGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGG 1889

Qy       1668 TGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGC 1727
Db       1890 TGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGC 1949

Qy       1728 GCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACAC 1787
Db       1950 GCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACAC 2009

Qy       1788 AGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAA 1847
Db       2010 AGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAA 2069

Qy       1848 ATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTG 1907
Db       2070 ATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTG 2129

Qy       1908 TTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAG 1967
Db       2130 TTACAGCTGTACAGCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAG 2189

Qy       1968 GAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGA 2027
Db       2190 GAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGA 2249

Qy       2028 TAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAA 2087
Db       2250 TAACCCCATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAA 2309

Qy       2088 CTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCAT 2147
```

```
Db      2310 CTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCAT 2369
Qy      2148 TGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGA 2207
Db      2370 TGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGA 2429
Qy      2208 CAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAA 2267
Db      2430 CAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAA 2489
Qy      2268 GAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAA 2327
Db      2490 GAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAA 2549
Qy      2328 GCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGA 2387
Db      2550 GCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGA 2609
Qy      2388 AGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGC 2447
Db      2610 AGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGC 2669
Qy      2448 TGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCAT 2507
Db      2670 TGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCAT 2729
Qy      2508 CAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCG 2567
Db      2730 CAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCG 2789
Qy      2568 GACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCAT 2627
Db      2790 GACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCAT 2849
Qy      2628 GTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTT 2687
Db      2850 GTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTT 2909
Qy      2688 TGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTA 2747
Db      2910 TGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTA 2969
Qy      2748 CTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAA 2807
Db      2970 CTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAA 3029
Qy      2808 TACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCA 2867
Db      3030 TACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCA 3089
Qy      2868 AGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGA 2927
Db      3090 AGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGA 3149
Qy      2928 CAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAA 2987
Db      3150 CAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAA 3209
Qy      2988 ATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGA 3047
Db      3210 ATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGA 3269
Qy      3048 GACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAA 3107
Db      3270 GACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAA 3329
Qy      3108 GTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAA 3167
Db      3330 GTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAA 3389
Qy      3168 GGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCC 3227
Db      3390 GGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCC 3449
Qy      3228 TCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATC 3287
Db      3450 TCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATC 3509
Qy      3288 CCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAA 3347
Db      3510 CCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAA 3569
Qy      3348 CAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGC 3407
Db      3570 CAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGC 3629
Qy      3408 TGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAA 3467
Db      3630 TGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAA 3689
Qy      3468 CATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAA 3527
Db      3690 CATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAA 3749
Qy      3528 TATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGATTGCCTT 3587
Db      3750 TATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGATTGCCTT 3809
Qy      3588 CTCTCAGCACAGCAATTTTATGGACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTT 3647
Db      3810 CTCTCAGCACAGCAATTTTATGGACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTT 3869
Qy      3648 CCTCTCTTTGCTCCTGGTGGCTGCTGTGGTTTGGAAGATCAAACAAAGTTGTTGGGCCTC 3707
Db      3870 CCTCTCTTTGCTCCTGGTGGCTGCTGTGGTTTGGAAGATCAAACAAAGTTGTTGGGCCTC 3929
Qy      3708 CAGACGTAGAGAGCAACTTCTTCGAGAGATGCAACAGATGGCCAGCCGTCCCTTTGCCTC 3767
Db      3930 CAGACGTAGAGAGCAACTTCTTCGAGAGATGCAACAGATGGCCAGCCGTCCCTTTGCCTC 3989
Qy      3768 TGTAAATGTCGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTATTGGGGGGAGTATAAA 3827
Db      3990 TGTAAATGTCGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTATTGGGGGGAGTATAAA 4049
Qy      3828 GACTGTTCCCAAACCCATTGCACTGGAGCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTC 3887
Db      4050 GACTGTTCCCAAACCCATTGCACTGGAGCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTC 4109
Qy      3888 TGTGTTTGTGAGGCTCCCTCGAGGCCTGGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCT 3947
Db      4110 TGTGTTTGTGAGGCTCCCTCGAGGCCTGGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCT 4169
Qy      3948 TGCTGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAA 4007
Db      4170 TGCTGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAA 4229
Qy      4008 GTCAGGAGCCGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGCCTGGGACCTGCATCTG 4067
Db      4230 GTCAGGAGCCGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGCCTGGGACCTGCATCTG 4289
Qy      4068 A 4068
Db      4290 A 4290
```

```
RESULT 5
AAZ52274
ID   AAZ52274 standard; cDNA; 3597 BP.
XX
AC   AAZ52274;
XX
DT   18-JUL-2000  (first entry)
XX
DE   Human soluble attractin-1 cDNA.
XX
```

```
KW   Human; soluble attractin-1; immune response; macrophage; monocyte;
KW   T cell; immunostimulant; immunosuppressed patient; cancer;
KW   immunodeficiency syndrome; transplant; autoimmune disease; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .3597
FT                   /*tag=  a
FT                   /product= "Soluble attractin-1"
XX
PN   WO200015651-A1.
XX
PD   23-MAR-2000.
XX
PF   14-SEP-1999;   99WO-US020948.
XX
PR   14-SEP-1998;   98US-0100137P.
XX
PA   (DAND ) DANA FARBER CANCER INST INC.
XX
PI   Duke-Cohan JS,  Schlossman SF;
XX
DR   WPI; 2000-271373/23.
DR   P-PSDB; AAY70689.
XX
PT   Isolated nucleic acids encoding human attractin polypeptides useful for
PT   enhancing immune responses.
XX
PS   Claim 3; Fig 8; 120pp; English.
XX
CC   The patent discloses four forms of human attractin polypeptides which
CC   enhance immune response by promoting macrophage and monocyte spreading in
CC   the presence of T cells. These include soluble attractin-1 and -2 and
CC   membrane attractin-1 and -2. These various forms of attractin are encoded
CC   by alternatively spliced mRNA molecule transcribed from a single gene.
CC   The present sequence is a cDNA encoding soluble attractin-1 obtained from
CC   activated T cell and human foetal liver libraries. This sequence can be
CC   used to enhance immune response in immunosuppressed patients such as
CC   those undergoing chemo- and radio-therapy treatment for cancer or those
CC   suffering from common variable immunodeficiency syndrome. The proteins
CC   may also be used to screen modulators (agonists and antagonists) of
CC   immune responses which may also be used to regulate immune reactions.
CC   Attractin antibodies can be used to inhibit immune response in transplant
CC   recipients or patients afflicted with autoimmune disease
XX
SQ   Sequence 3597 BP; 951 A; 822 C; 898 G; 926 T; 0 U; 0 Other;

  Query Match             87.4%;  Score 3556.4;  DB 3;  Length 3597;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 3579;  Conservative    0;  Mismatches    1;  Indels    2;  Gaps    2;

Qy          1 ATGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 60

Qy         61 GCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGC 120

Qy        121 ACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGC 180

Qy        181 CGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATAC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATAC 240

Qy        241 AAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTC 300

Qy        301 AATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATT 360

Qy        361 TATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAG 420

Qy        421 ACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCT 480

Qy        481 GCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGC 540

Qy        541 TCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCT 600

Qy        601 GAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCT 660

Qy        661 CATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAG 720

Qy        721 GGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATAT 780

Qy        781 TCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGG 840

Qy        841 GTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTT 900

Qy        901 GCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGT 960

Qy        961 CATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACT 1020

Qy       1021 GGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTG 1080

Qy       1081 ACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTG 1140

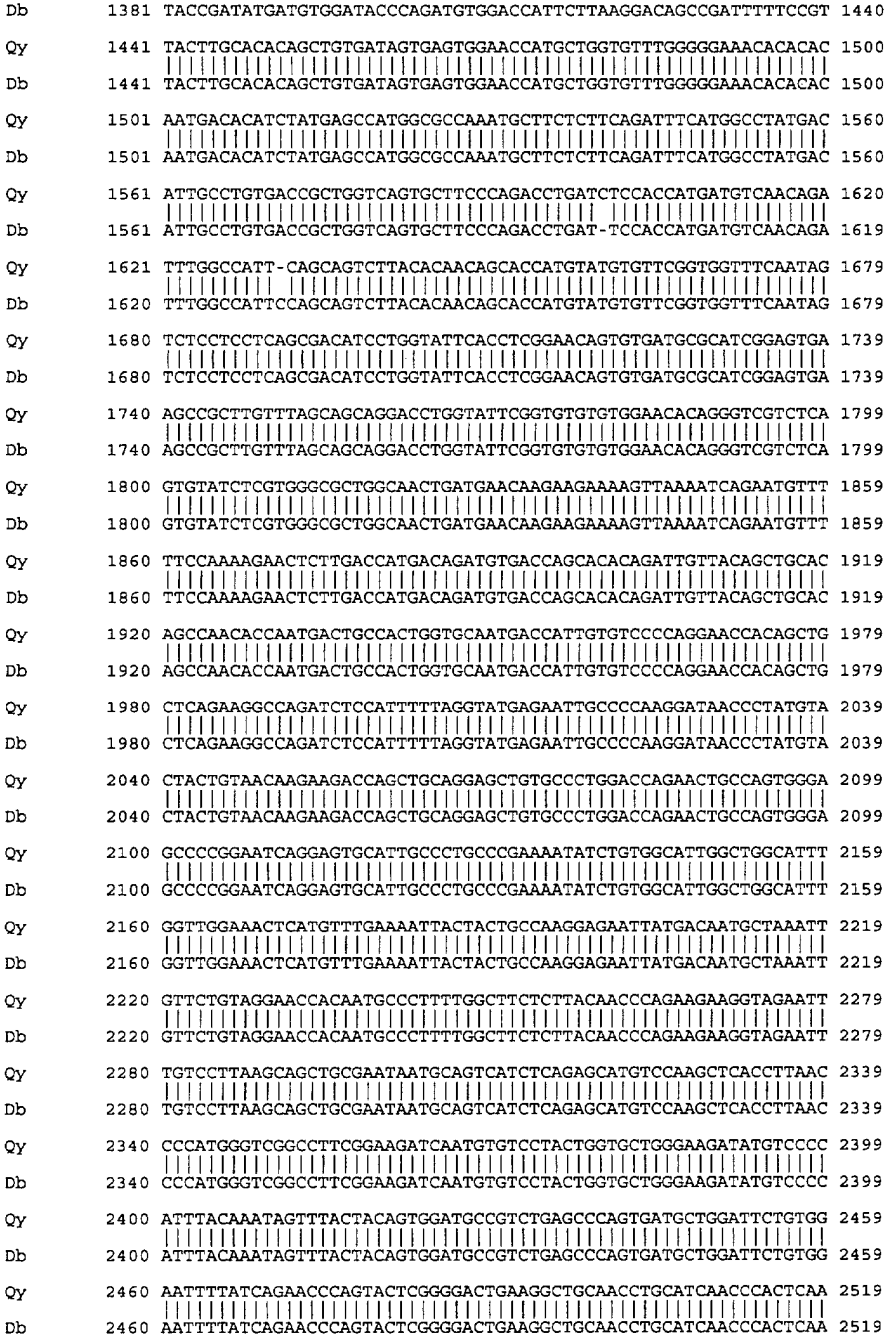
Qy       1141 AAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATA 1200

Qy       1201 AGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGT 1260

Qy       1261 GCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTA 1320

Qy       1321 TACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTC 1380

Qy       1381 TACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGT 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```


Db 1381 TACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGT 1440
Qy 1441 TACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACAC 1500
Db 1441 TACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACAC 1500
Qy 1501 AATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGAC 1560
Db 1501 AATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGAC 1560
Qy 1561 ATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGA 1620
Db 1561 ATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGAT-TCCACCATGATGTCAACAGA 1619
Qy 1621 TTTGGCCATT-CAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAG 1679
Db 1620 TTTGGCCATTCCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAG 1679
Qy 1680 TCTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGA 1739
Db 1680 TCTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGA 1739
Qy 1740 AGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCA 1799
Db 1740 AGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCA 1799
Qy 1800 GTGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTT 1859
Db 1800 GTGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTT 1859
Qy 1860 TTCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCAC 1919
Db 1860 TTCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCAC 1919
Qy 1920 AGCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTG 1979
Db 1920 AGCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTG 1979
Qy 1980 CTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTA 2039
Db 1980 CTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTA 2039
Qy 2040 CTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGA 2099
Db 2040 CTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGA 2099
Qy 2100 GCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTT 2159
Db 2100 GCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTT 2159
Qy 2160 GGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATT 2219
Db 2160 GGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATT 2219
Qy 2220 GTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATT 2279
Db 2220 GTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATT 2279
Qy 2280 TGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAAC 2339
Db 2280 TGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAAC 2339
Qy 2340 CCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCC 2399
Db 2340 CCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCC 2399
Qy 2400 ATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGG 2459
Db 2400 ATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGG 2459
Qy 2460 AATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAA 2519
Db 2460 AATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAA 2519
Qy 2520 TGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGC 2579
Db 2520 TGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGC 2579
Qy 2580 CTTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAA 2639
Db 2580 CTTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAA 2639
Qy 2640 CATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTAT 2699
Db 2640 CATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTAT 2699
Qy 2700 GGAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAG 2759
Db 2700 GGAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAG 2759
Qy 2760 TCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGG 2819
Db 2760 TCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGG 2819
Qy 2820 GAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGG 2879
Db 2820 GAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGG 2879
Qy 2880 AAATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAA 2939
Db 2880 AAATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAA 2939
Qy 2940 CTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCA 2999
Db 2940 CTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCA 2999
Qy 3000 GAGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATC 3059
Db 3000 GAGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATC 3059
Qy 3060 TGGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCA 3119
Db 3060 TGGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCA 3119
Qy 3120 CGCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGG 3179
Db 3120 CGCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGG 3179
Qy 3180 GGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAAC 3239
Db 3180 GGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAAC 3239
Qy 3240 ATGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGA 3299
Db 3240 ATGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGA 3299
Qy 3300 TCGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGA 3359
Db 3300 TCGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGA 3359
Qy 3360 CATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTC 3419
Db 3360 CATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTC 3419
Qy 3420 AGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTA 3479
Db 3420 AGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTA 3479
Qy 3480 CAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTT 3539
Db 3480 CAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTT 3539
Qy 3540 TGTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGAT 3581
Db 3540 TGTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGGT 3581

```
RESULT 6
AAS72657
ID   AAS72657 standard; cDNA; 3597 BP.
XX
AC   AAS72657;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #8461.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG08470.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 8461; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 3597 BP; 951 A; 822 C; 898 G; 926 T; 0 U; 0 Other;

  Query Match             87.4%;  Score 3556.4;  DB 5;  Length 3597;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 3579;  Conservative    0;  Mismatches    1;  Indels    2;  Gaps    2;

Qy          1 ATGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 60

Qy         61 GCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGC 120

Qy        121 ACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGC 180

Qy        181 CGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATAC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATAC 240

Qy        241 AAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTC 300

Qy        301 AATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATT 360

Qy        361 TATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAG 420

Qy        421 ACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCT 480

Qy        481 GCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGC 540

Qy        541 TCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCT 600

Qy        601 GAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCT 660

Qy        661 CATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAG 720

Qy        721 GGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATAT 780

Qy        781 TCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGG 840

Qy        841 GTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTT 900

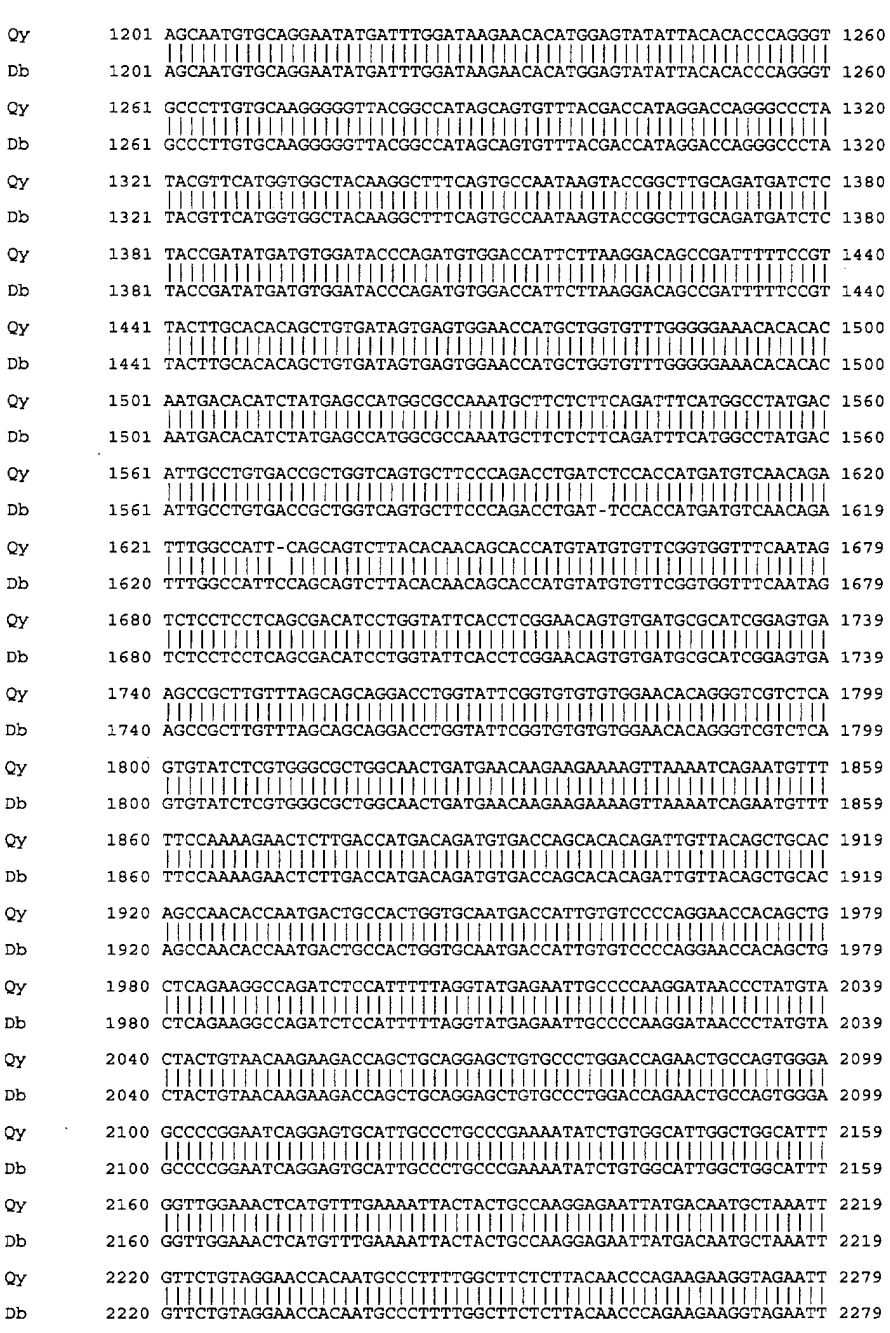
Qy        901 GCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGT 960

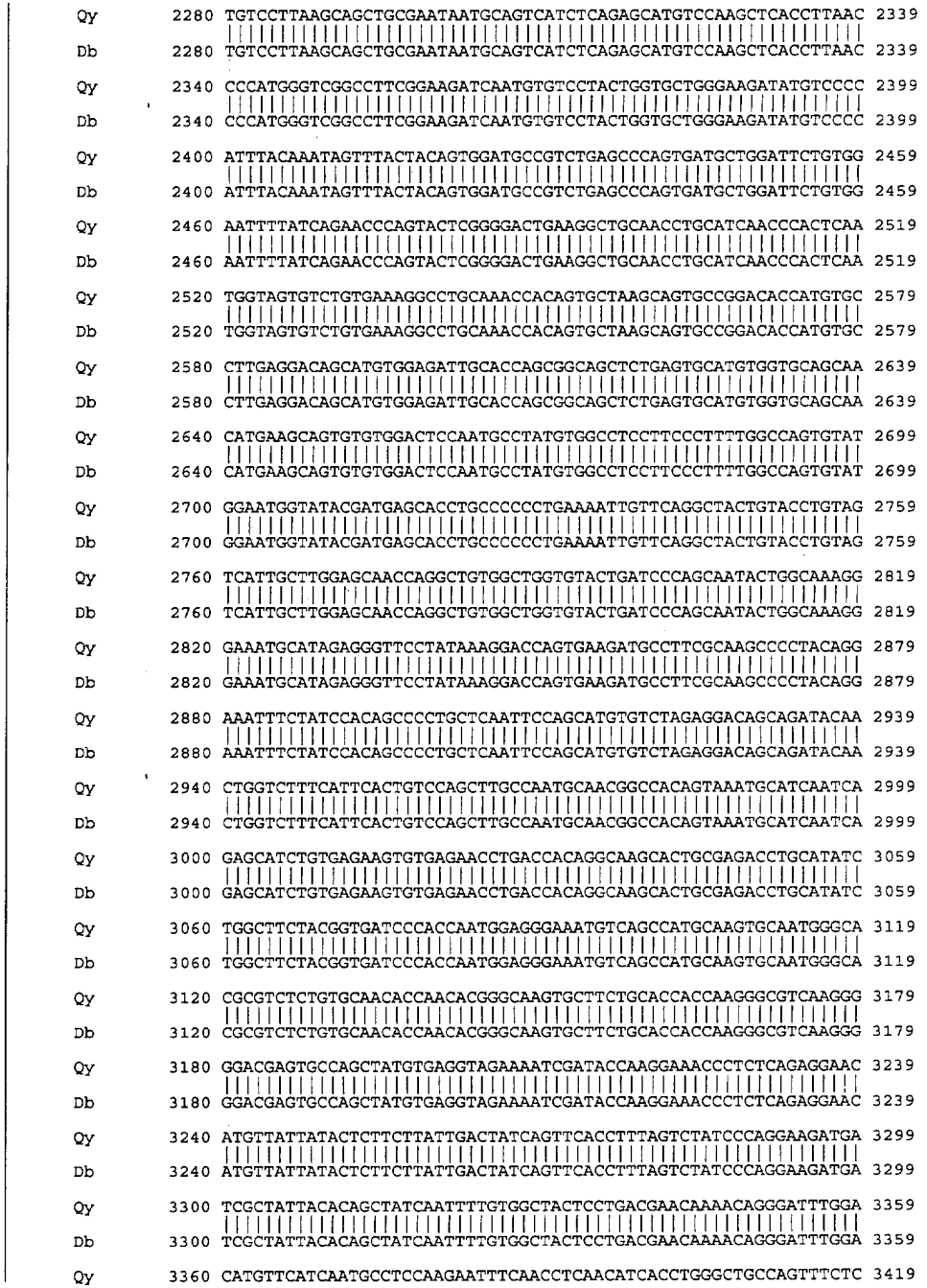
Qy        961 CATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACT 1020

Qy       1021 GGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTG 1080

Qy       1081 ACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTG 1140

Qy       1141 AAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATA 1200
```


Qy 1201 AGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGT 1260
Db 1201 AGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGT 1260
Qy 1261 GCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTA 1320
Db 1261 GCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTA 1320
Qy 1321 TACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTC 1380
Db 1321 TACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTC 1380
Qy 1381 TACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGT 1440
Db 1381 TACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGT 1440
Qy 1441 TACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACAC 1500
Db 1441 TACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACAC 1500
Qy 1501 AATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGAC 1560
Db 1501 AATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGAC 1560
Qy 1561 ATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGA 1620
Db 1561 ATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGAT-TCCACCATGATGTCAACAGA 1619
Qy 1621 TTTGGCCATT-CAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAG 1679
Db 1620 TTTGGCCATTCCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAG 1679
Qy 1680 TCTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGA 1739
Db 1680 TCTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGA 1739
Qy 1740 AGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCA 1799
Db 1740 AGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCA 1799
Qy 1800 GTGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTT 1859
Db 1800 GTGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTT 1859
Qy 1860 TTCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCAC 1919
Db 1860 TTCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCAC 1919
Qy 1920 AGCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTG 1979
Db 1920 AGCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTG 1979
Qy 1980 CTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTA 2039
Db 1980 CTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTA 2039
Qy 2040 CTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGA 2099
Db 2040 CTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGA 2099
Qy 2100 GCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTT 2159
Db 2100 GCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTT 2159
Qy 2160 GGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATT 2219
Db 2160 GGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATT 2219
Qy 2220 GTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATT 2279
Db 2220 GTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATT 2279


Qy 2280 TGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAAC 2339
Db 2280 TGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAAC 2339
Qy 2340 CCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCC 2399
Db 2340 CCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCC 2399
Qy 2400 ATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGG 2459
Db 2400 ATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGG 2459
Qy 2460 AATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAA 2519
Db 2460 AATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAA 2519
Qy 2520 TGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGC 2579
Db 2520 TGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGC 2579
Qy 2580 CTTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAA 2639
Db 2580 CTTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAA 2639
Qy 2640 CATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTAT 2699
Db 2640 CATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTAT 2699
Qy 2700 GGAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAG 2759
Db 2700 GGAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAG 2759
Qy 2760 TCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGG 2819
Db 2760 TCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGG 2819
Qy 2820 GAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGG 2879
Db 2820 GAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGG 2879
Qy 2880 AAATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAA 2939
Db 2880 AAATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAA 2939
Qy 2940 CTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCA 2999
Db 2940 CTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCA 2999
Qy 3000 GAGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATC 3059
Db 3000 GAGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATC 3059
Qy 3060 TGGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCA 3119
Db 3060 TGGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCA 3119
Qy 3120 CGCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGG 3179
Db 3120 CGCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGG 3179
Qy 3180 GGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAAC 3239
Db 3180 GGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAAC 3239
Qy 3240 ATGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGA 3299
Db 3240 ATGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGA 3299
Qy 3300 TCGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGA 3359
Db 3300 TCGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGA 3359
Qy 3360 CATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTC 3419

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3360 CATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTC 3419

Qy       3420 AGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTA 3479
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3420 AGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTA 3479

Qy       3480 CAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTT 3539
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3480 CAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTT 3539

Qy       3540 TGTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGAT 3581
                |||||||||||||||||||||||||||||||||||||||| |
Db       3540 TGTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGGT 3581
```

RESULT 7
ADD70997
ID ADD70997 standard; DNA; 3597 BP.
XX
AC ADD70997;
XX
DT 15-JAN-2004 (first entry)
XX
DE Human attractin gene SEQ ID NO:1.
XX
KW liver cancer; chronic hepatitis; cirrhosis; liver disease; hepatotropic;
KW cytostatic; gene therapy; human; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2003061564-A2.
XX
PD 31-JUL-2003.
XX
PF 20-DEC-2002; 2002WO-US040718.
XX
PR 21-DEC-2001; 2001US-0341815P.
PR 31-DEC-2001; 2001US-0343185P.
XX
PA (GENE-) GENE LOGIC INC.
PA (LGBI-) LG BIOMEDICAL INST.
XX
PI Koh SS, Liu Q, Chung H, Zeng W, Lee B, Yeramilli S, Song SY;
XX
DR WPI; 2003-663343/62.
XX
PT Diagnosing liver cancer cells, useful for treating liver cancer
PT associated with chronic hepatitis or cirrhosis comprises detecting the
PT level of expression in a tissue sample of one or more genes associated
PT with cancerous liver tissues.
XX
PS Claim 1; SEQ ID NO 1; 176pp; English.
XX
CC The present invention describes a method for diagnosing liver cancer
CC cells comprising detecting the level of expression in a tissue sample of
CC one or more genes given in the specification (see ADD70997 to ADD71105),
CC where differential expression of the genes is indicative of liver cancer.
CC Also described: (1) detecting the progression of liver cancer in a
CC patient; (2) monitoring the treatment of a patient with liver cancer; (3)
CC treating a patient with liver cancer; (4) typing a liver disease in a
CC patient; (5) detecting the presence or progression of liver cancer in a
CC patient with chronic hepatitis or cirrhosis; (6) differentiating liver
CC cancer related to chronic hepatitis from liver cancer related to
CC cirrhosis; (7) screening for an agent capable of modulating the onset or
CC progression of liver cancer; (8) a composition comprising at least two
CC oligonucleotides comprising a sequence that specifically hybridises to
CC any of the genes; (9) a solid support comprising the at least two
CC oligonucleotides; (10) a computer system comprising a database containing
CC information identifying the level in liver tissue of a set of genes; (11)
CC a method for using the computer system to present information identifying
CC the expression level in tissue or cell of any of the genes; and (12) a
CC therapeutic agent for slowing or halting the progression of liver cancer.
CC The methods are useful for treating liver cancer associated with chronic
CC hepatitis or cirrhosis. The present sequence represents a specifically
CC claimed human gene sequence which is used in the exemplification of the
CC present invention.
XX
SQ Sequence 3597 BP; 951 A; 822 C; 898 G; 926 T; 0 U; 0 Other;

Query Match 87.4%; Score 3556.4; DB 9; Length 3597;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3579; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

```
Qy          1 ATGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 60

Qy         61 GCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGC 120

Qy        121 ACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGC 180

Qy        181 CGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATAC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATAC 240

Qy        241 AAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTC 300

Qy        301 AATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATT 360

Qy        361 TATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAG 420

Qy        421 ACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCT 480

Qy        481 GCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGC 540

Qy        541 TCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCT 600

Qy        601 GAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCT 660

Qy        661 CATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAG 720

Qy        721 GGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATAT 780

Qy        781 TCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGG 840

Qy        841 GTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTT 900

Qy        901 GCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        901 GCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGT 960
Qy        961 CATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACT 1020
Db        961 CATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACT 1020
Qy       1021 GGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTG 1080
Db       1021 GGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTG 1080
Qy       1081 ACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTG 1140
Db       1081 ACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTG 1140
Qy       1141 AAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATA 1200
Db       1141 AAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATA 1200
Qy       1201 AGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGT 1260
Db       1201 AGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGT 1260
Qy       1261 GCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTA 1320
Db       1261 GCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTA 1320
Qy       1321 TACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTC 1380
Db       1321 TACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTC 1380
Qy       1381 TACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGT 1440
Db       1381 TACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGT 1440
Qy       1441 TACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACAC 1500
Db       1441 TACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACAC 1500
Qy       1501 AATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGAC 1560
Db       1501 AATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGAC 1560
Qy       1561 ATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGA 1620
Db       1561 ATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGAT-TCCACCATGATGTCAACAGA 1619
Qy       1621 TTTGGCCATT-CAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAG 1679
Db       1620 TTTGGCCATTCCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAG 1679
Qy       1680 TCTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGA 1739
Db       1680 TCTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGA 1739
Qy       1740 AGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCA 1799
Db       1740 AGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCA 1799
Qy       1800 GTGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTT 1859
Db       1800 GTGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTT 1859
Qy       1860 TTCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCAC 1919
Db       1860 TTCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCAC 1919
Qy       1920 AGCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTG 1979
Db       1920 AGCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTG 1979
Qy       1980 CTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTA 2039
Db       1980 CTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTA 2039
```

```
Qy       2040 CTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGA 2099
Db       2040 CTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGA 2099
Qy       2100 GCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTT 2159
Db       2100 GCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTT 2159
Qy       2160 GGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATT 2219
Db       2160 GGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATT 2219
Qy       2220 GTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATT 2279
Db       2220 GTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATT 2279
Qy       2280 TGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAAC 2339
Db       2280 TGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAAC 2339
Qy       2340 CCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCC 2399
Db       2340 CCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCC 2399
Qy       2400 ATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGG 2459
Db       2400 ATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGG 2459
Qy       2460 AATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAA 2519
Db       2460 AATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAA 2519
Qy       2520 TGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGC 2579
Db       2520 TGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGC 2579
Qy       2580 CTTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAA 2639
Db       2580 CTTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAA 2639
Qy       2640 CATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTAT 2699
Db       2640 CATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTAT 2699
Qy       2700 GGAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAG 2759
Db       2700 GGAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAG 2759
Qy       2760 TCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGG 2819
Db       2760 TCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGG 2819
Qy       2820 GAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGG 2879
Db       2820 GAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGG 2879
Qy       2880 AAATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAA 2939
Db       2880 AAATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAA 2939
Qy       2940 CTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCA 2999
Db       2940 CTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCA 2999
Qy       3000 GAGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATC 3059
Db       3000 GAGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATC 3059
Qy       3060 TGGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCA 3119
Db       3060 TGGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCA 3119
```

```
Qy       3120 CGCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGG 3179
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3120 CGCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGG 3179

Qy       3180 GGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAAC 3239
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3180 GGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAAC 3239

Qy       3240 ATGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGA 3299
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3240 ATGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGA 3299

Qy       3300 TCGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGA 3359
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3300 TCGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGA 3359

Qy       3360 CATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTC 3419
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3360 CATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTC 3419

Qy       3420 AGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTA 3479
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3420 AGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTA 3479

Qy       3480 CAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTT 3539
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3480 CAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTT 3539

Qy       3540 TGTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGAT 3581
             |||||||||||||||||||||||||||||||||||||||||| |
Db       3540 TGTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGGT 3581
```

RESULT 8
AAZ91921
ID AAZ91921 standard; cDNA; 4072 BP.
XX
AC AAZ91921;
XX
DT 08-JUN-2000 (first entry)
XX
DE Human mahogany protein coding sequence #3.
XX
KW Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;
KW weight regulation; cell therapy; body weight disorder; cachexia;
KW anorexia; hyperpigmentation; increased metabolic rate disorder;
KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.
XX
OS Homo sapiens.
XX
PN WO200005373-A2.
XX
PD 03-FEB-2000.
XX
PF 21-JUL-1999; 99WO-US016484.
XX
PR 21-JUL-1998; 98US-0093630P.
PR 20-OCT-1998; 98US-0104978P.
PR 05-FEB-1999; 99US-00245041.
XX
PA (MILL-) MILLENIUM PHARM INC.
XX
PI Moore K, Nagle DL;
XX
DR WPI; 2000-195103/17.
DR P-PSDB; AAY81808.
XX
PT New human and murine mahogany genes, useful, e.g. for diagnosis and
PT treatment of body weight disorders.
XX
PS Claim 1; Fig 19a; 188pp; English.
XX
CC This sequence represents a human mahogany gene of the invention. The
CC mahogany genes are used: (i) to produce recombinant mahogany (mg)
CC proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming
CC therapeutics; (iii) as a source of diagnostic probes and primers for
CC detecting expression of mg genes or mutations, regulatory defects, in
CC this gene, or for isolation of related sequences; and (iv) in (cell-
CC based) gene therapy. (II) are used to raise specific antibodies (Ab); to
CC identify other (extra)cellular products involved in weight regulation,
CC and to screen for agents that disrupt interaction between (II) and other
CC macromolecules. The Ab are used to detect abnormal levels (or function)
CC of (II) (for diagnosis, prognosis or monitoring of treatment); to
CC evaluate (II)-expressing cells intended for cell therapy, and as
CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the
CC mg polypeptide) are used to identify agents (A) that modulate mg
CC activity. (A) are potentially useful for the treatment of body weight
CC disorders, particularly obesity, cachexia or anorexia, or other
CC conditions associated with the mg gene such as hyperpigmentation,
CC hyperphagia and disorders that result in increased metabolic rate
XX
SQ Sequence 4072 BP; 1087 A; 940 C; 1045 G; 1000 T; 0 U; 0 Other;

Query Match 86.0%; Score 3499.8; DB 3; Length 4072;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 3525; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```
Qy         15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74
              || |||    |  ||   | ||  |  |  |  | | |   |||| ||   || || ||
Db        234 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAA 293

Qy         75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134
              |   |  |  | ||  |||||||||||||||||||||||||||||||||||||||||||||
Db        294 GGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 353

Qy        135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        354 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 413

Qy        195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 473

Qy        255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 533

Qy        315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 593

Qy        375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 653

Qy        435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        654 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 713

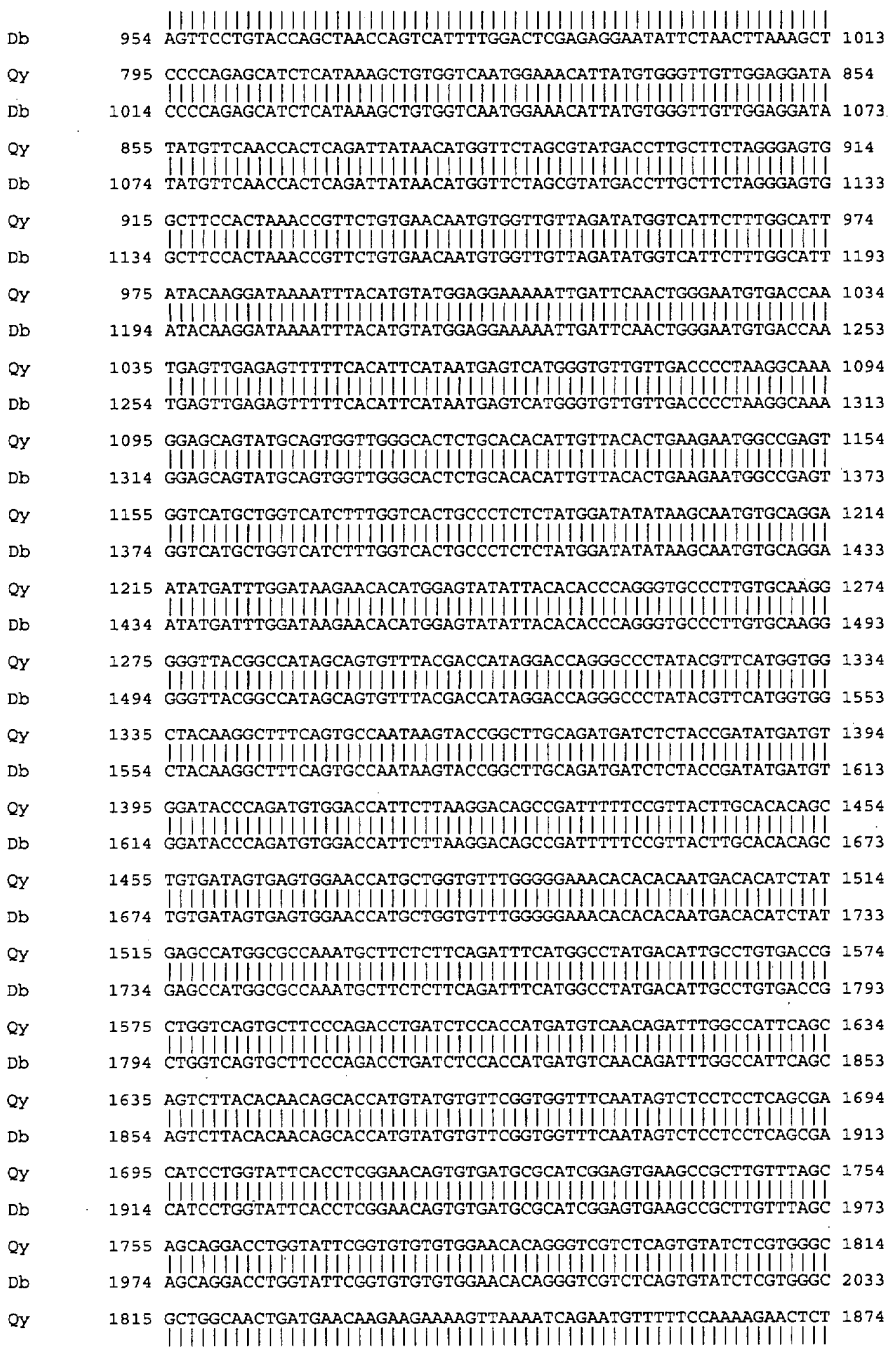
Qy        495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        714 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 773

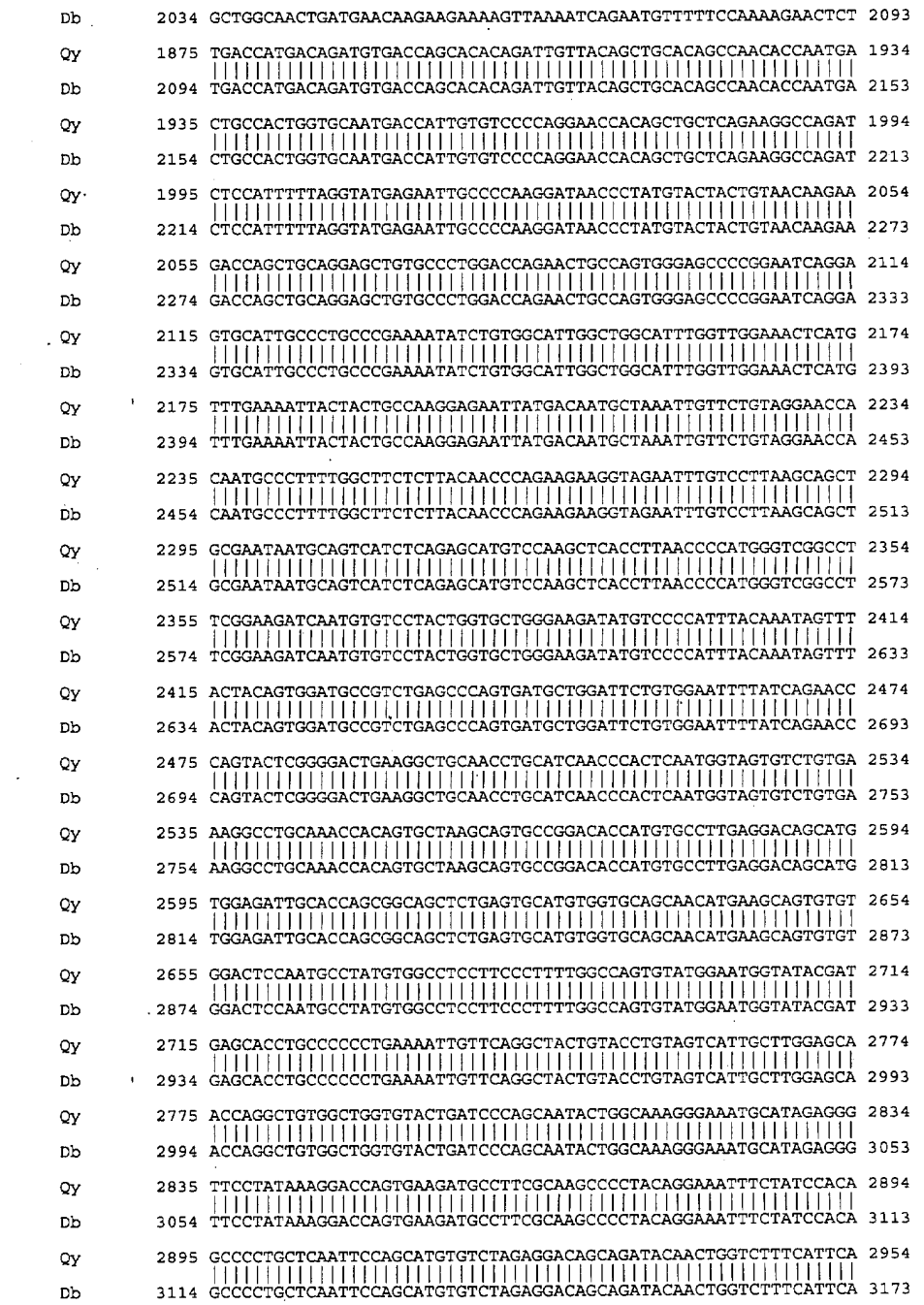
Qy        555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        774 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 833

Qy        615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        834 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 893

Qy        675 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        894 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 953

Qy        735 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 794
```


Db 954 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 1013
Qy 795 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 854
Db 1014 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 1073
Qy 855 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 914
Db 1074 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 1133
Qy 915 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 974
Db 1134 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 1193
Qy 975 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1034
Db 1194 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1253
Qy 1035 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1094
Db 1254 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1313
Qy 1095 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1154
Db 1314 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1373
Qy 1155 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1214
Db 1374 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1433
Qy 1215 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1274
Db 1434 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1493
Qy 1275 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1334
Db 1494 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1553
Qy 1335 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1394
Db 1554 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1613
Qy 1395 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1454
Db 1614 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1673
Qy 1455 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1514
Db 1674 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1733
Qy 1515 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1574
Db 1734 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1793
Qy 1575 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGC 1634
Db 1794 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGC 1853
Qy 1635 AGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGA 1694
Db 1854 AGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGA 1913
Qy 1695 CATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGC 1754
Db 1914 CATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGC 1973
Qy 1755 AGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGC 1814
Db 1974 AGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGC 2033
Qy 1815 GCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCT 1874


Db 2034 GCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCT 2093
Qy 1875 TGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGA 1934
Db 2094 TGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGA 2153
Qy 1935 CTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGAT 1994
Db 2154 CTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGAT 2213
Qy 1995 CTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAA 2054
Db 2214 CTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAA 2273
Qy 2055 GACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGA 2114
Db 2274 GACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGA 2333
Qy 2115 GTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATG 2174
Db 2334 GTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATG 2393
Qy 2175 TTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCA 2234
Db 2394 TTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCA 2453
Qy 2235 CAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCT 2294
Db 2454 CAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCT 2513
Qy 2295 GCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCT 2354
Db 2514 GCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCT 2573
Qy 2355 TCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTT 2414
Db 2574 TCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTT 2633
Qy 2415 ACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACC 2474
Db 2634 ACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACC 2693
Qy 2475 CAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGA 2534
Db 2694 CAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGA 2753
Qy 2535 AAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATG 2594
Db 2754 AAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATG 2813
Qy 2595 TGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGT 2654
Db 2814 TGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGT 2873
Qy 2655 GGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGAT 2714
Db 2874 GGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGAT 2933
Qy 2715 GAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCA 2774
Db 2934 GAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCA 2993
Qy 2775 ACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGG 2834
Db 2994 ACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGG 3053
Qy 2835 TTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACA 2894
Db 3054 TTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACA 3113
Qy 2895 GCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCA 2954
Db 3114 GCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCA 3173

```
Qy       2955 CTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAA 3014
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3174 CTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAA 3233

Qy       3015 GTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGA 3074
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3234 GTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGA 3293

Qy       3075 TCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAA 3134
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3294 TCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAA 3353

Qy       3135 CACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCT 3194
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3354 CACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCT 3413

Qy       3195 ATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCT 3254
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3414 ATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCT 3473

Qy       3255 TCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGC 3314
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3474 TCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGC 3533

Qy       3315 TATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGC 3374
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3534 TATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGC 3593

Qy       3375 CTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGC 3434
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3594 CTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGC 3653

Qy       3435 TGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTC 3494
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3654 TGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTC 3713

Qy       3495 TAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAA 3554
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3714 TAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAA 3773

Qy       3555 TTTCACCTGGCCCATCAAAATTCAGAT 3581
             |||||||||||||||||||||||||||
Db       3774 TTTCACCTGGCCCATCAAAATTCAGGT 3800
```

RESULT 9
AAZ52277
ID AAZ52277 standard; cDNA; 3819 BP.
XX
AC AAZ52277;
XX
DT 18-JUL-2000 (first entry)
XX
DE Human soluble attractin-2 cDNA.
XX
KW Human; soluble attractin-2; immune response; macrophage; monocyte;
KW T cell; immunostimulant; immunosuppressed patient; cancer;
KW immunodeficiency syndrome; transplant; autoimmune disease; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1. .3819
FT /*tag= a
FT /product= "Soluble attractin-2"
XX
PN WO200015651-A1.
XX
PD 23-MAR-2000.
XX
PF 14-SEP-1999; 99WO-US020948.
XX
PR 14-SEP-1998; 98US-0100137P.
XX
PA (DAND) DANA FARBER CANCER INST INC.
XX
PI Duke-Cohan JS, Schlossman SF;
XX
DR WPI; 2000-271373/23.
DR P-PSDB; AAY70692.
XX
PT Isolated nucleic acids encoding human attractin polypeptides useful for
PT enhancing immune responses.
XX
PS Claim 3; Fig 12; 120pp; English.
XX
CC The patent discloses four forms of human attractin polypeptides which
CC enhance immune response by promoting macrophage and monocyte spreading in
CC the presence of T cells. These include soluble attractin-1 and -2 and
CC membrane attractin-1 and -2. These various forms of attractin are encoded
CC by alternatively spliced mRNA molecule transcribed from a single gene.
CC The present sequence is a cDNA encoding human soluble attractin-2.
CC Attractin-2 differs from attractin-1 in having a 74 amino acid insert in
CC the N-terminal. This sequence can be used to enhance immune response in
CC immunosuppressed patients such as those undergoing chemo- and radio-
CC therapy treatment for cancer or those suffering from common variable
CC immunodeficiency syndrome. The proteins may also be used to screen
CC modulators (agonists and antagonists) of immune response which may also
CC be used to regulate immune reactions. Attractin antibodies can be used to
CC inhibit immune response in transplant recipients or patients afflicted
CC with autoimmune disease
XX
SQ Sequence 3819 BP; 969 A; 899 C; 991 G; 960 T; 0 U; 0 Other;

Query Match 85.9%; Score 3492.4; DB 3; Length 3819;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 3508; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

```
Qy         48 GGCGGCGACGGCAGCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGTGTCAACGGCGGTCG 107
              | |||  | || ||  ||.|| || |  | |  | || ||||||||||||||||||
Db        270 GTCGGGCTCAGCCGCAGCCGAGGCCAAGGAATGTGACCGGCCCTGTGTCAACGGCGGTCG 329

Qy        108 CTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCA 167
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        330 CTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCA 389

Qy        168 GCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGG 227
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        390 GCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGG 449

Qy        228 AAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAAT 287
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        450 AAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAAT 509

Qy        288 GAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGA 347
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        510 GAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGA 569

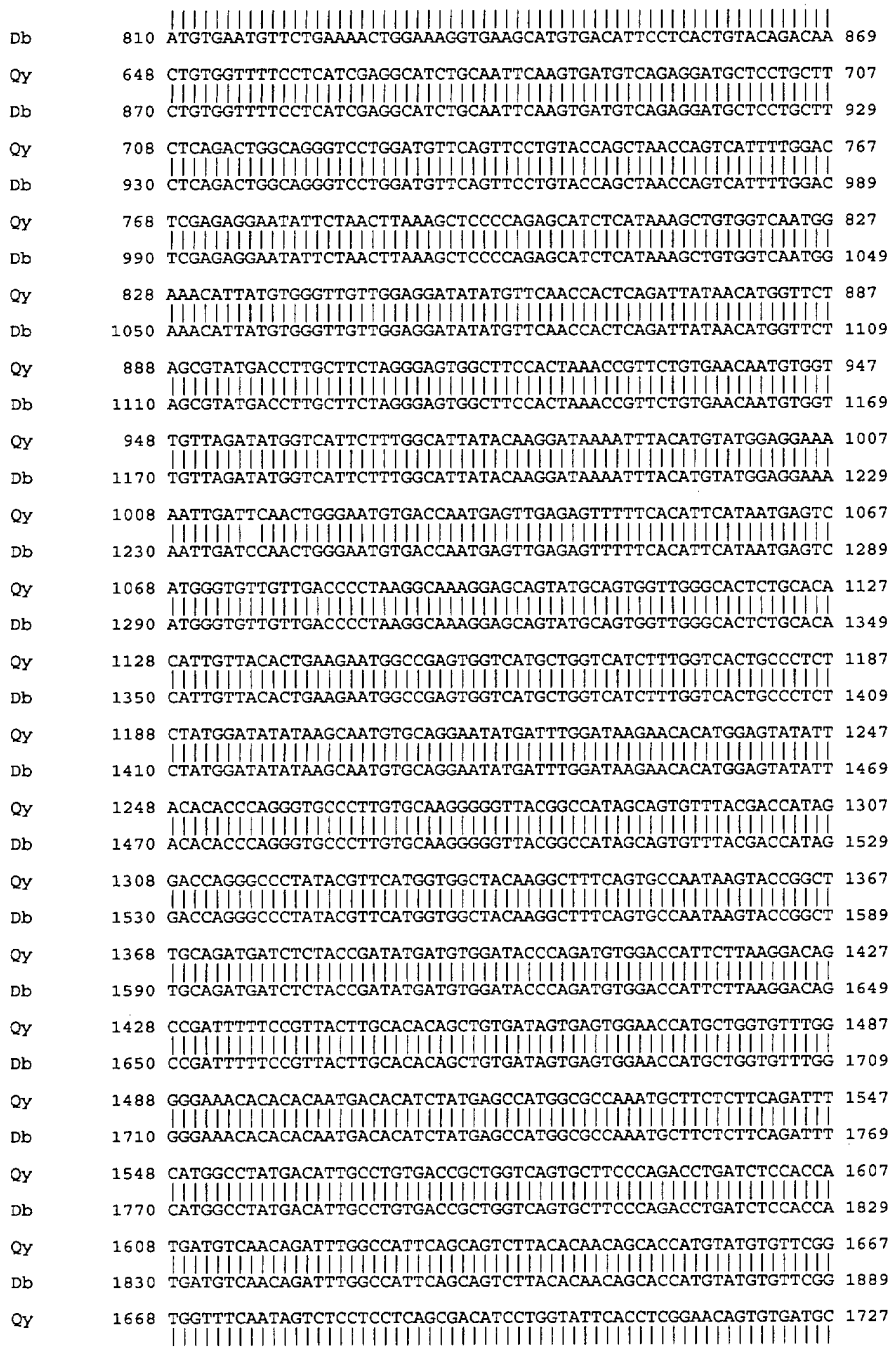
Qy        348 TGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAG 407
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        570 TGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAG 629

Qy        408 AGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTT 467
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        630 AGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTT 689

Qy        468 TTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTG 527
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        690 TTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTG 749

Qy        528 TCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGA 587
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        750 TCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGAAACTGTTGA 809

Qy        588 ATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA 647
```


Db 810 ATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA 869
Qy 648 CTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTT 707
Db 870 CTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTT 929
Qy 708 CTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGAC 767
Db 930 CTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGAC 989
Qy 768 TCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGG 827
Db 990 TCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGG 1049
Qy 828 AAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCT 887
Db 1050 AAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCT 1109
Qy 888 AGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGT 947
Db 1110 AGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGT 1169
Qy 948 TGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAA 1007
Db 1170 TGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAA 1229
Qy 1008 AATTGATTCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTC 1067
Db 1230 AATTGATCCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTC 1289
Qy 1068 ATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACA 1127
Db 1290 ATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACA 1349
Qy 1128 CATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCT 1187
Db 1350 CATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCT 1409
Qy 1188 CTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATT 1247
Db 1410 CTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATT 1469
Qy 1248 ACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAG 1307
Db 1470 ACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAG 1529
Qy 1308 GACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCT 1367
Db 1530 GACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCT 1589
Qy 1368 TGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAG 1427
Db 1590 TGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAG 1649
Qy 1428 CCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGG 1487
Db 1650 CCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGG 1709
Qy 1488 GGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTT 1547
Db 1710 GGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTT 1769
Qy 1548 CATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCA 1607
Db 1770 CATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCA 1829
Qy 1608 TGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGG 1667
Db 1830 TGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGG 1889
Qy 1668 TGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGC 1727
Db 1890 TGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGC 1949
Qy 1728 GCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACAC 1787
Db 1950 GCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACAC 2009
Qy 1788 AGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAA 1847
Db 2010 AGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAA 2069
Qy 1848 ATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTG 1907
Db 2070 ATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTG 2129
Qy 1908 TTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAG 1967
Db 2130 TTACAGCTGTACAGCCAACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAG 2189
Qy 1968 GAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGA 2027
Db 2190 GAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGA 2249
Qy 2028 TAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAA 2087
Db 2250 TAACCCCATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAA 2309
Qy 2088 CTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCAT 2147
Db 2310 CTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCAT 2369
Qy 2148 TGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGA 2207
Db 2370 TGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGA 2429
Qy 2208 CAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAA 2267
Db 2430 CAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAA 2489
Qy 2268 GAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAA 2327
Db 2490 GAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAA 2549
Qy 2328 GCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGA 2387
Db 2550 GCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGA 2609
Qy 2388 AGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGC 2447
Db 2610 AGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGC 2669
Qy 2448 TGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCAT 2507
Db 2670 TGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCAT 2729
Qy 2508 CAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCG 2567
Db 2730 CAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCG 2789
Qy 2568 GACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCAT 2627
Db 2790 GACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCAT 2849
Qy 2628 GTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTT 2687
Db 2850 GTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTT 2909
Qy 2688 TGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTA 2747
Db 2910 TGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTA 2969
Qy 2748 CTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAA 2807
Db 2970 CTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAA 3029

```
Qy       2808 TACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCA 2867
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3030 TACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCA 3089

Qy       2868 AGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGA 2927
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3090 AGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGA 3149

Qy       2928 CAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAA 2987
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3150 CAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAA 3209

Qy       2988 ATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGA 3047
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3210 ATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGA 3269

Qy       3048 GACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAA 3107
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3270 GACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAA 3329

Qy       3108 GTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAA 3167
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3330 GTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAA 3389

Qy       3168 GGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCC 3227
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3390 GGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCC 3449

Qy       3228 TCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATC 3287
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3450 TCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATC 3509

Qy       3288 CCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAA 3347
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3510 CCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAA 3569

Qy       3348 CAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGC 3407
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3570 CAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGC 3629

Qy       3408 TGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAA 3467
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3630 TGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAA 3689

Qy       3468 CATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAA 3527
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3690 CATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAA 3749

Qy       3528 TATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGAT 3581
             ||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       3750 TATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGGT 3803
```

```
RESULT 10
ADB53249
ID   ADB53249 standard; DNA; 8739 BP.
XX
AC   ADB53249;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3791.
XX
KW   toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KW   toxicity marker; toxicity progression; drug screening;
KW   primary rat hepatocyte toxicity modelling; gene; ds.
XX
OS   Rattus norvegicus.
XX
PN   WO2003065993-A2.
XX
PD   14-AUG-2003.
XX
PF   04-FEB-2003; 2003WO-US003482.
XX
PR   04-FEB-2002; 2002US-0353171P.
PR   13-MAR-2002; 2002US-0363534P.
PR   08-APR-2002; 2002US-0370248P.
PR   10-APR-2002; 2002US-0371134P.
PR   10-APR-2002; 2002US-0371135P.
PR   10-APR-2002; 2002US-0371150P.
PR   11-APR-2002; 2002US-0371413P.
PR   19-APR-2002; 2002US-0373601P.
PR   19-APR-2002; 2002US-0373602P.
PR   22-APR-2002; 2002US-0374139P.
PR   08-MAY-2002; 2002US-0378370P.
PR   09-MAY-2002; 2002US-0378652P.
PR   09-MAY-2002; 2002US-0378653P.
PR   09-MAY-2002; 2002US-0378665P.
PR   09-JUL-2002; 2002US-0394230P.
PR   09-JUL-2002; 2002US-0394253P.
PR   04-SEP-2002; 2002US-0407688P.
PR   28-JAN-2003; 2003US-0442900P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Mendrick D,  Porter M,  Johnson K,  Higgs B,  Castle A,  Orr M;
PI   Elashoff M;
XX
DR   WPI; 2003-731472/69.
XX
PT   Determining if a compound induces a toxic effect on a tissue or cell, for
PT   identifying hepatotoxic compounds, comprises comparing a gene expression
PT   profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT   mean values.
XX
PS   Claim 44; SEQ ID NO 3791; 874pp; English.
XX
CC   The present invention describes a method for determining whether a
CC   compound induces a toxic effect on a tissue or cell. The method comprises
CC   preparing a gene expression profile of a tissue or cell sample exposed to
CC   the compound, and comparing the gene expression profile to a database
CC   comprising data or information on the Tox mean and non-Tox mean value.
CC   The method is useful for predicting or identifying at least one toxic
CC   effect, particularly hepatotoxicity, of a test or unknown compound. The
CC   genes listed in the specification are useful as diagnostic or toxicity
CC   markers for the prediction or identification of the physiological state
CC   of tissue or cell sample that has been exposed to a compound, or to
CC   identify or predict the toxic effects of a compound or an agent. These
CC   may also be used as markers for monitoring toxicity progression or for
CC   drug screening. The present sequence represents a primary rat hepatocyte
CC   toxicity modelling related gene sequence from the present invention.
XX
SQ   Sequence 8739 BP; 2288 A; 2115 C; 2109 G; 2227 T; 0 U; 0 Other;

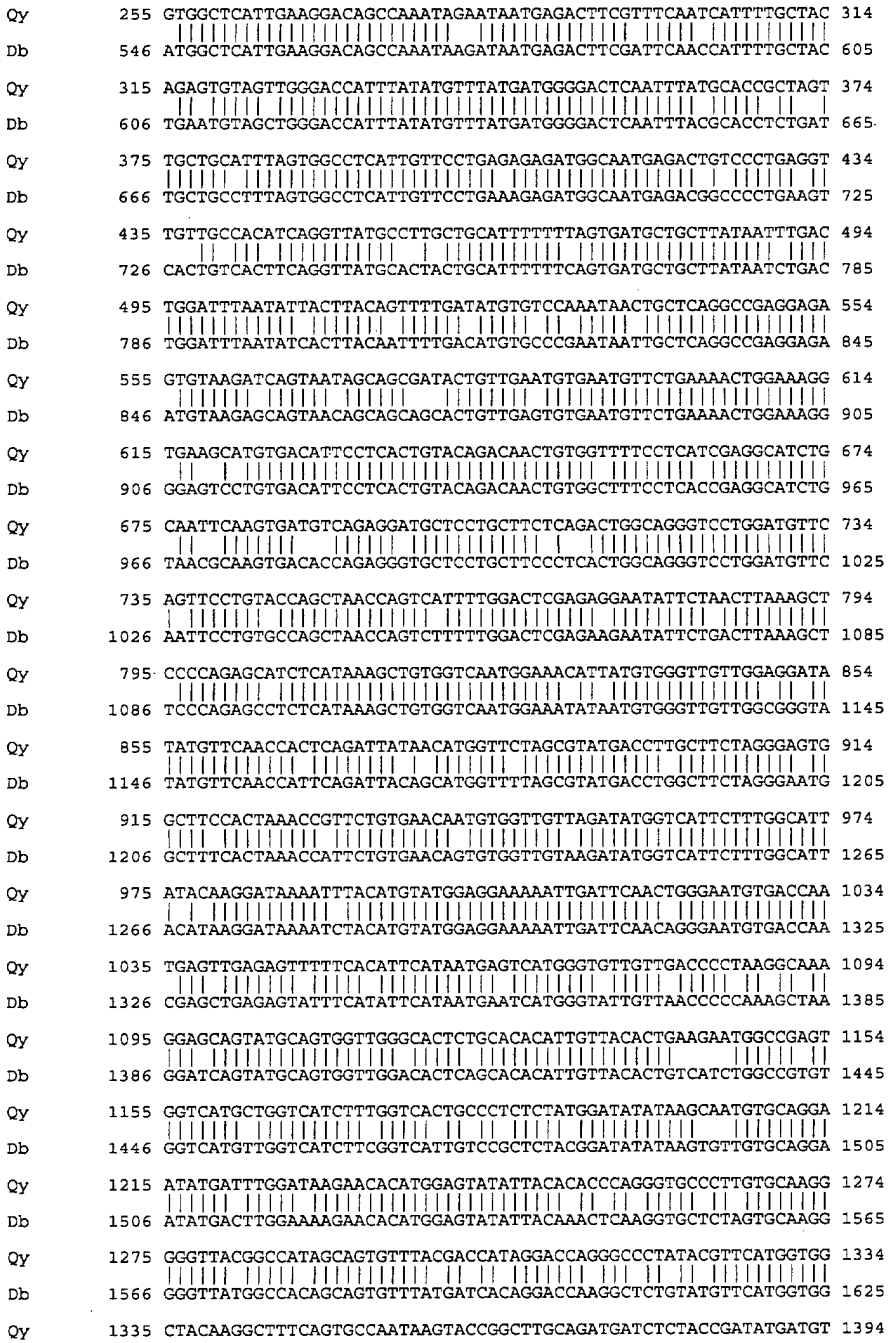
  Query Match           82.4%;  Score 3353.2;  DB 9;  Length 8739;
  Best Local Similarity  89.2%;  Pred. No. 0;
  Matches 3616;  Conservative    0;  Mismatches  438;  Indels    0;  Gaps    0;

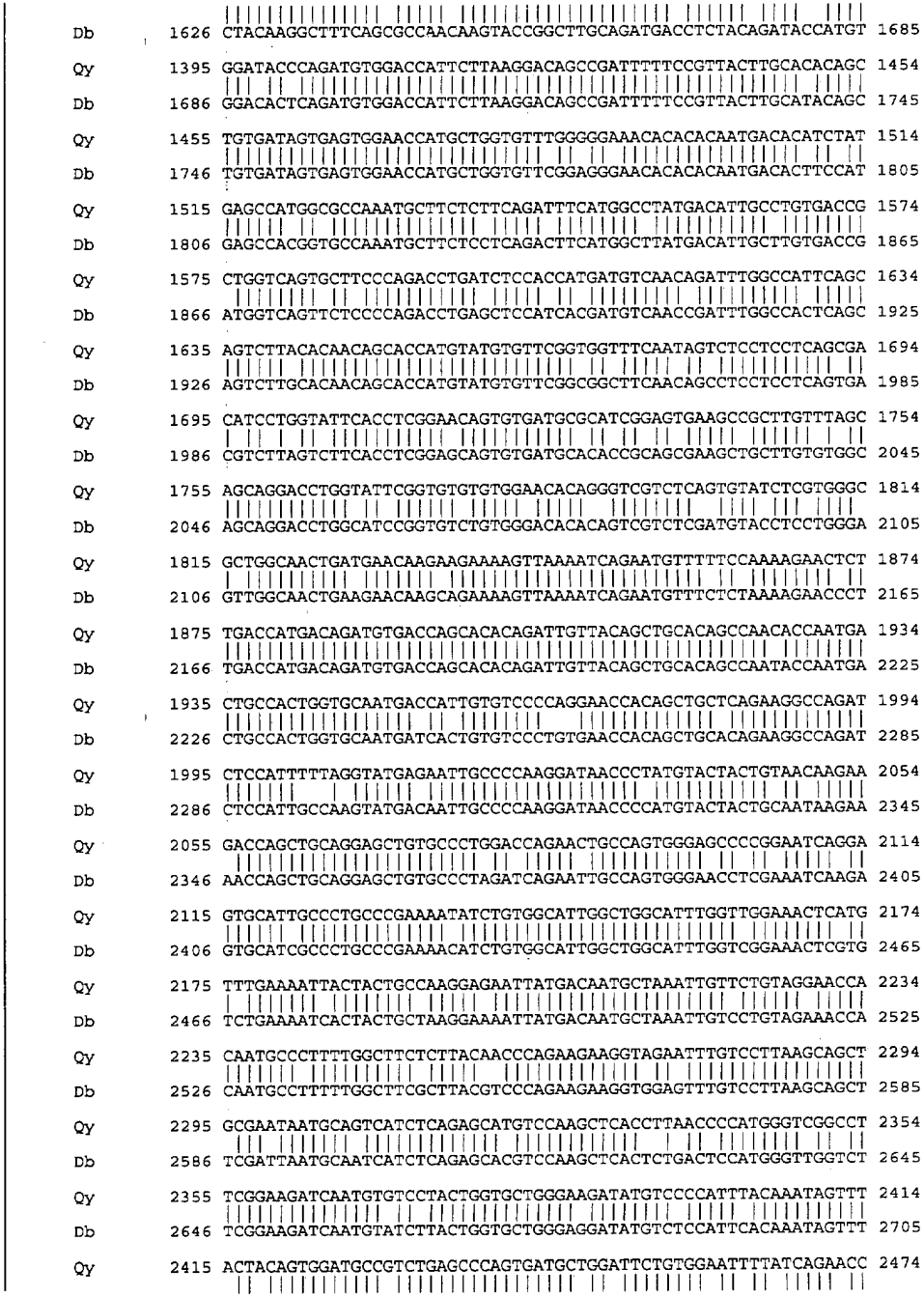
Qy         15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74
              || |||     |  ||   | ||  |  |      | | |   |||| ||     | || ||
Db        306 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCTGTGTCAGGCTCGGCCGCAGCTGAGGCCAA 365

Qy         75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134
              |    | |  | ||  |||||||||||||| ||||||||||||||||||||||||||||||
Db        366 GGAATGTGACCGGCCGTGTGTCAACGGCGGCCGCTGCAACCCTGGCACCGGCCAGTGCGT 425

Qy        135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194
              ||||||| | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 CTGCCCCACGGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 485

Qy        195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254
              |||||||||||| ||||| |||||||||||||| |||||||||| |||||||| |||||
Db        486 TGGATCTTCTGGATTTGTAACAGATGGACCTGGGAATTATAAATATAAAACGAAATGCAC 545
```





```
Db       2706 GCTGCAGTGGATGCCATCTGAGCCCAGTGATGCCGGCTTCTGTGGGATCTTGTCAGAGCC 2765
Qy       2475 CAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGA 2534
Db       2766 TAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGAAGTGTTTGTGA 2825
Qy       2535 AAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATG 2594
Db       2826 AAGGCCTGCAAACCACAGTGCCAAGCAGTGCCGGACACCCTGTGCCCTGCGGACAGCATG 2885
Qy       2595 TGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGT 2654
Db       2886 TGGTGAGTGCACCAGCAGCAGCTCCGAATGCATGTGGTGCAGCAACATGAAGCAGTGCGT 2945
Qy       2655 GGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGAT 2714
Db       2946 AGACTCCAATGCCTACGTAGCCTCCTTTCCTTTCGGCCAGTGTATGGAGTGGTATACAAT 3005
Qy       2715 GAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCA 2774
Db       3006 GAGCAGCTGCCCACCTGAAAATTGCTCTGGCTACTGTACCTGCAGCCATTGCTTGGAGCA 3065
Qy       2775 ACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGG 2834
Db       3066 GCCAGGCTGTGGTTGGTGTACTGATCCTAGCAATACTGGGAAAGGCAAATGTATTGAGGG 3125
Qy       2835 TTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACA 2894
Db       3126 AAGCTATAAAGGACCCGTGAAGATGCCTTCACATGCCTCTACAGGAAATGTGTACCCACA 3185
Qy       2895 GCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCA 2954
Db       3186 GCCCCTTCTGAACTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTTATTCA 3245
Qy       2955 CTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAA 3014
Db       3246 CTGTCCAGCTTGCCAGTGCAATGGACACAGCAAGTGCATCAACCAGAGCATCTGTGAGAA 3305
Qy       3015 GTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGA 3074
Db       3306 GTGTGAGGACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTATGGTGA 3365
Qy       3075 TCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAA 3134
Db       3366 CCCGACTAATGGAGGCAAATGTCAGCCATGCAAGTGCAACGGGCACGCATCACTGTGCAA 3425
Qy       3135 CACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCT 3194
Db       3426 CACCAACACTGGCAAGTGCTTCTGTACCACCAAGGGCGTCAAGGGAGAGGAGTGCCAGCT 3485
Qy       3195 ATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCT 3254
Db       3486 ATGTGAGGTAGAAAATCGATACCAGGGAAACCCTCTCAAAGGAACATGTTACTATACCCT 3545
Qy       3255 TCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGC 3314
Db       3546 TCTCATTGACTATCAGTTCACTTTTAGCCTGTCCCAGGAAGACGACCGCTATTACACAGC 3605
Qy       3315 TATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGC 3374
Db       3606 CATCAATTTTGTGGCTACTCCTGATGAACAAAACAGGGATTTGGACATGTTCATCAATGC 3665
Qy       3375 CTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGC 3434
Db       3666 CTCCAAAAACTTCAACCTCAACATCACCTGGGCCACCAGTTTCCCAGCTGGAACCCAGAC 3725
Qy       3435 TGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTC 3494
Db       3726 TGGAGAAGAGGTGCCTGTTGTTTCAAAAACCAACATCAAGGAATACAAAGATAGCTTCTC 3785
Qy       3495 TAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAA 3554
Db       3786 TAATGAGAAATTTGATTTTCGCAACCATCCAAACATTACTTTCTTTGTTTATGTCAGTAA 3845
```

```
Qy       3555 TTTCACCTGGCCCATCAAAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCT 3614
Db       3846 TTTCACTTGGCCCATCAAAATTCAAATTGCCTTCTCCCAGCACAGCAATTTTATGGACCT 3905
Qy       3615 GGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGT 3674
Db       3906 GGTGCAGTTCTTCGTGACTTTCTTCAGCTGTTTCCTCTCTCTGCTTCTGGTGGCGGCAGT 3965
Qy       3675 GGTTTGGAAGATCAAACAAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGA 3734
Db       3966 GGTCTGGAAGATCAAGCAGAGCTGTTGGGCATCCAGGAGGAGAGAGCAACTTCTTCGGGA 4025
Qy       3735 GATGCAACAGATGGCCAGCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGA 3794
Db       4026 GATGCAACAGATGGCCAGCCGCCCTTTTGCTTCTGTAAATGTTGCCTTGGAAACAGATGA 4085
Qy       3795 GGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGA 3854
Db       4086 AGAACCTCCTGATCTCATTGGGGGAAGTATAAAGACTGTTCCTAAGCCCATTGCCCTGGA 4145
Qy       3855 GCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCT 3914
Db       4146 GCCCTGCTTTGGTAACAAAGCTGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGACT 4205
Qy       3915 GGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACAT 3974
Db       4206 GGGAGGAATCCCTCCTCCTGGTCAGTCAGGTCTTGCTGTGGCCAGTGCCCTGGTGGACAT 4265
Qy       3975 TTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCA 4034
Db       4266 TTCTCAGCAGATGCCAATAGTGTACAAGGAGAAGTCAGGAGCTGTAAGGAACCGGAAGCA 4325
Qy       4035 GCAGCCCCCTGCACAGCCTGGGACCTGCATCTGA 4068
Db       4326 GCAGCCCCCTGCACAGCCTGGAACCTGCATTTGA 4359
```

RESULT 11
AAZ91916
ID AAZ91916 standard; cDNA; 8827 BP.
XX
AC AAZ91916;
XX
DT 08-JUN-2000 (first entry)
XX
DE Wild type (C57BL/6J) murine mahogany protein coding sequence.
XX
KW Mahogany gene; mouse; mg gene; regulatory defect; gene therapy; obesity;
KW weight regulation; cell therapy; body weight disorder; cachexia;
KW anorexia; hyperpigmentation; increased metabolic rate disorder;
KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.
XX
OS Mus sp.
XX
PN WO200005373-A2.
XX
PD 03-FEB-2000.
XX
PF 21-JUL-1999; 99WO-US016484.
XX
PR 21-JUL-1998; 98US-0093630P.
PR 20-OCT-1998; 98US-0104978P.
PR 05-FEB-1999; 99US-00245041.
XX
PA (MILL-) MILLENIUM PHARM INC.
XX
PI Moore K, Nagle DL;
XX
DR WPI; 2000-195103/17.
DR P-PSDB; AAY81803.
XX
PT New human and murine mahogany genes, useful, e.g. for diagnosis and

PT treatment of body weight disorders.
XX
PS Claim 1; Fig 2a; 188pp; English.
XX
CC This sequence represents a murine mahogany gene of the invention. The
CC mahogany genes are used: (i) to produce recombinant mahogany (mg)
CC proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming
CC therapeutics; (iii) as a source of diagnostic probes and primers for
CC detecting expression of mg genes or mutations, regulatory defects, in
CC this gene, or for isolation of related sequences; and (iv) in (cell-
CC based) gene therapy. (II) are used to raise specific antibodies (Ab); to
CC identify other (extra)cellular products involved in weight regulation,
CC and to screen for agents that disrupt interaction between (II) and other
CC macromolecules. The Ab are used to detect abnormal levels (or function)
CC of (II) (for diagnosis, prognosis or monitoring of treatment); to
CC evaluate (II)-expressing cells intended for cell therapy, and as
CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the
CC mg polypeptide) are used to identify agents (A) that modulate mg
CC activity. (A) are potentially useful for the treatment of body weight
CC disorders, particularly obesity, cachexia or anorexia, or other
CC conditions associated with the mg gene such as hyperpigmentation,
CC hyperphagia and disorders that result in increased metabolic rate
XX
SQ Sequence 8827 BP; 2304 A; 2077 C; 2162 G; 2284 T; 0 U; 0 Other;

Query Match 81.7%; Score 3324.8; DB 3; Length 8827;
Best Local Similarity 91.1%; Pred. No. 0;
Matches 3533; Conservative 0; Mismatches 347; Indels 0; Gaps 0;

```
Qy       189 ACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAA 248
Db       677 ACTAACTGGCTCTTCTGGATTTGTAACAGATGGACCTGGGAATTATAAATATAAGACGAA 736

Qy       249 GTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTT 308
Db       737 GTGCACATGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGCTTCAACCATTT 796

Qy       309 TGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACC 368
Db       797 TGCTACAGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATCTACGCACC 856

Qy       369 GCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCC 428
Db       857 TCTGATTGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCTCC 916

Qy       429 TGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAA 488
Db       917 TGAGGTCACTGTCACTTCAGGTTATGCACTGCTGCATTTTTTCAGTGATGCTGCTTATAA 976

Qy       489 TTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCG 548
Db       977 TCTGACTGGATTTAATATCACTTACAATTTTGACATGTGTCCGAATAATTGCTCAGGCCG 1036

Qy       549 AGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTG 608
Db       1037 AGGAGAGTGTAAGAGCAGTAACAGCAGCAGCGCTGTTGAGTGTGAATGTTCTGAAAACTG 1096

Qy       609 GAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGG 668
Db       1097 GAAAGGGGAGTCGTGTGACATTCCTCACTGTACAGACAACTGTGGCTTTCCTCACCGAGG 1156

Qy       669 CATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGG 728
Db       1157 CATCTGTAATGCAAGCGATACCAGAGGGTGCTCCTGCTTTCCTCACTGGCAGGGTCCTGG 1216

Qy       729 ATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTT 788
Db       1217 ATGTTCAATTCCTGTGCCAGCTAACCAGTCTTTTTGGACTCGAGAAGAATATTCTGATTT 1276

Qy       789 AAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGG 848
Db       1277 AAAGCTTCCCAGAGCCTCTCATAAAGCTGTGGTCAATGGAAATATAATGTGGGTTGTTGG 1336

Qy       849 AGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAG 908
Db       1337 CGGATATATGTTCAACCATTCAGATTACAGCATGGTTTTAGCGTATGACCTGACTTCTAG 1396

Qy       909 GGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTT 968
Db       1397 GGAATGGCTTCCACTAAACCATTCTGTGAACAGTGTGGTTGTAAGATATGGTCATTCTTT 1456

Qy       969 GGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGT 1028
Db       1457 GGCATTACATAAGGATAAAATCTACATGTATGGAGGAAAAATTGATTCAACAGGGAACGT 1516

Qy       1029 GACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAA 1088
Db       1517 GACCAATGAGCTGAGAGTATTTCATATTCATAATGAATCATGGGTATTGTTAACTCCGAA 1576

Qy       1089 GGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGG 1148
Db       1577 AGCTAAGGATCAGTATGCAGTGGTTGGACACTCAGCACACATTGTTACACTGGCATCTGG 1636

Qy       1149 CCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGT 1208
Db       1637 CCGTGTGGTCATGTTGGTCATCTTCGGTCATTGCCCACTCTATGGATATATAAGCGTTGT 1696

Qy       1209 GCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGT 1268
Db       1697 GCAGGAATATGACTTGGAAAAGAACACATGGAGTATATTACATACTCAGGGTGCTCTTGT 1756

Qy       1269 GCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCA 1328
Db       1757 GCAAGGGGGTTATGGCCACAGTAGTGTTTATGATGACAGGACCAAGGCTCTGTACGTTCA 1816

Qy       1329 TGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATA 1388
Db       1817 TGGTGGCTACAAGGCTTTCAGCGCCAACAAATACCGGCTTGCAGATGACCTCTACAGATA 1876

Qy       1389 TGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCA 1448
Db       1877 CGATGTGGATACTCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCA 1936

Qy       1449 CACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACAC 1508
Db       1937 TACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGAGGGAACACACACAATGACAC 1996

Qy       1509 ATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTG 1568
Db       1997 TTCCATGAGCCACGGTGCCAAATGCTTCTCCTCGGACTTCATGGCTTATGACATTGCTTG 2056

Qy       1569 TGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCA 1628
Db       2057 TGACCGATGGTCAGTGCTTCCCAGACCTGAGCTCCATCATGATGTCAACAGATTTGGCCA 2116

Qy       1629 TTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCT 1688
Db       2117 TTCAGCAGTCTTGTACAACAGCACCATGTATGTGTTCGGCGGCTTCAACAGCCTCCTCCT 2176

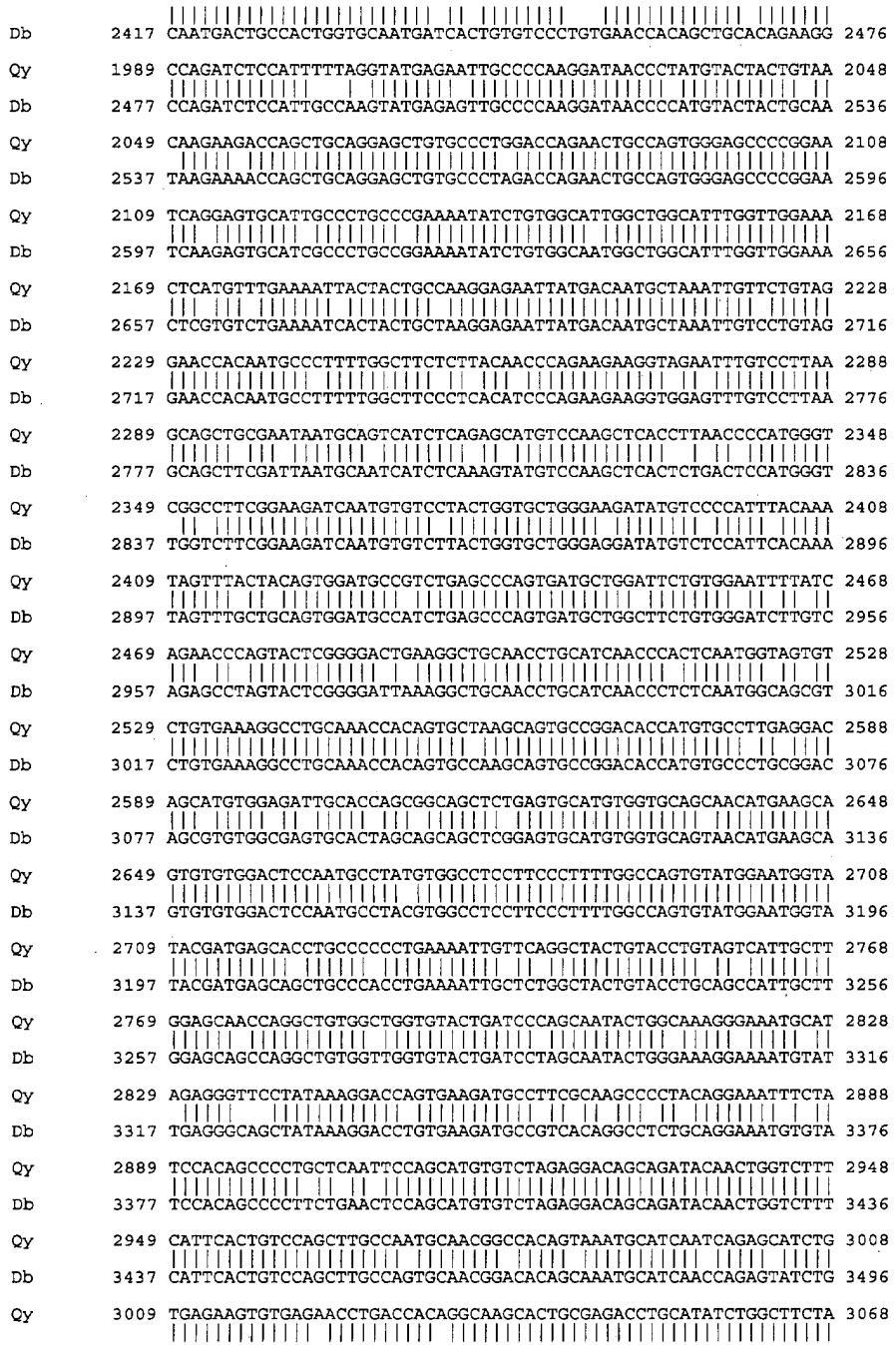
Qy       1689 CAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTG 1748
Db       2177 CAGTGACGTCTTGGTCTTTACCTCGGAGCAGTGCGATGCACACCGCAGTGAAGCTGCTTG 2236

Qy       1749 TTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTC 1808
Db       2237 TGTGGCAGCAGGACCTGGTATCCGGTGTCTGTGGGACACACAGTCGTCTCGATGTACCTC 2296

Qy       1809 GTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAG 1868
Db       2297 CTGGGAGTTGGCAACTGAAGAACAAGCAGAAAAGTTAAAATCAGAGTGTTTTTCTAAAAG 2356

Qy       1869 AACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACAC 1928
Db       2357 AACCCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAATAC 2416

Qy       1929 CAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGG 1988
```



RESULT 12
AAS72660

ID AAS72660 standard; cDNA; 3490 BP.
XX
AC AAS72660;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #8464.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG08473.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 8464; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 3490 BP; 931 A; 772 C; 858 G; 929 T; 0 U; 0 Other;

Query Match 69.1%; Score 2810.4; DB 5; Length 3490;
Best Local Similarity 94.3%; Pred. No. 0;
Matches 3236; Conservative 0; Mismatches 51; Indels 145; Gaps 25;

```
Qy        187 AGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACG 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 AGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACG 81

Qy        247 AAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCAT 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 AAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCAT 141

Qy        307 TTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCA 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 TTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCA 201

Qy        367 CCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTC 426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 CCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTC 261

Qy        427 CCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTAT 486
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 CCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTAT 321

Qy        487 AATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGC 546
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 AATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGC 381

Qy        547 CGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAAC 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        382 CGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAAC 441

Qy        607 TGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGA 666
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        442 TGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGA 501

Qy        667 GGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCT 726
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        502 GGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCT 561

Qy        727 GGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTT-GGACTCGAGAGGAATATTCTAA 785
              ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        562 GGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGGACTCGAGAGGAATATTCTAA 621

Qy        786 CTT--AAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTT 843
              |||  ||     ||||||||||||||||||||||||||||||||||||||||||||||
Db        622 CTTTAAAGCTCCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTT 681

Qy        844 GTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCT 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        682 GTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCT 741

Qy        904 TCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCAT 963
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        742 TCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCAT 801

Qy        964 TCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGG 1023
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        802 TCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGG 861

Qy       1024 AATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACC 1083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        862 AATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACC 921

Qy       1084 CCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAG 1143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        922 CCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAG 981

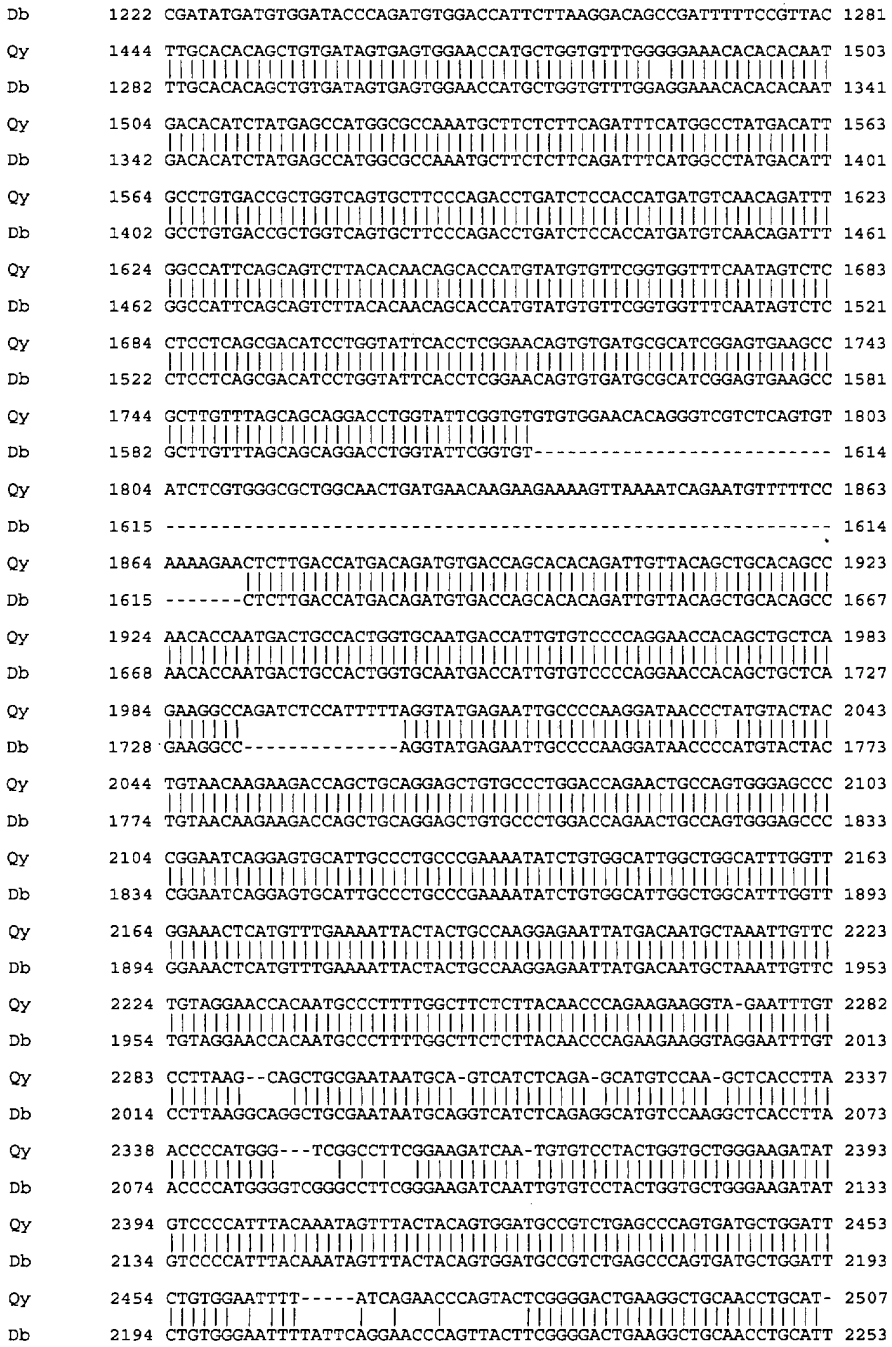
Qy       1144 AATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGC 1203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        982 AATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGC 1041

Qy       1204 AATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCC 1263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1042 AATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCC 1101

Qy       1264 CTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATAC 1323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1102 CTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATAC 1161

Qy       1324 GTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTAC 1383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1162 GTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTAC 1221

Qy       1384 CGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTAC 1443
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```


Db 1222 CGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTAC 1281
Qy 1444 TTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAAT 1503
Db 1282 TTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGAGGAAACACACACAAT 1341
Qy 1504 GACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATT 1563
Db 1342 GACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATT 1401
Qy 1564 GCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTT 1623
Db 1402 GCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTT 1461
Qy 1624 GGCCATTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTC 1683
Db 1462 GGCCATTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTC 1521
Qy 1684 CTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCC 1743
Db 1522 CTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCC 1581
Qy 1744 GCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGT 1803
Db 1582 GCTTGTTTAGCAGCAGGACCTGGTATTCGGTGT---------- 1614
Qy 1804 ATCTCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCC 1863
Db 1615 ---------- 1614
Qy 1864 AAAAGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCC 1923
Db 1615 -------CTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCC 1667
Qy 1924 AACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCA 1983
Db 1668 AACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCA 1727
Qy 1984 GAAGGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTAC 2043
Db 1728 GAAGGCC--------------AGGTATGAGAATTGCCCCAAGGATAACCCCATGTACTAC 1773
Qy 2044 TGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCC 2103
Db 1774 TGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCC 1833
Qy 2104 CGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTT 2163
Db 1834 CGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTT 1893
Qy 2164 GGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTC 2223
Db 1894 GGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTC 1953
Qy 2224 TGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTA-GAATTTGT 2282
Db 1954 TGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGGAATTTGT 2013
Qy 2283 CCTTAAG--CAGCTGCGAATAATGCA-GTCATCTCAGA-GCATGTCCAA-GCTCACCTTA 2337
Db 2014 CCTTAAGGCAGGCTGCGAATAATGCAGGTCATCTCAGAGGCATGTCCAAGGCTCACCTTA 2073
Qy 2338 ACCCCATGGG---TCGGCCTTCGGAAGATCAA-TGTGTCCTACTGGTGCTGGGAAGATAT 2393
Db 2074 ACCCCATGGGGTCGGGCCTTCGGGAAGATCAATTGTGTCCTACTGGTGCTGGGAAGATAT 2133
Qy 2394 GTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATT 2453
Db 2134 GTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATT 2193
Qy 2454 CTGTGGAATTTT-----ATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCAT- 2507
Db 2194 CTGTGGGAATTTTATTCAGGAACCCAGTTACTTCGGGGACTGAAGGCTGCAACCTGCATT 2253
Qy 2508 -CAACCCACTCAATGGTAGTGTCTGTGAAA--GGCCTGCAAACCACAGT-GCTAAGCAGT 2563
Db 2254 CAACCCACTTCAATGGTAGTGTCTGTGAAAAGGCCTGGCAAACCACAGTGGTTAAGCAGT 2313
Qy 2564 GCCGGACACCATGTGCC----TTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCT 2619
Db 2314 GCCGGACACCATGGTGCCCTTTGAGGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCT 2373
Qy 2620 GAGT-GCATGTGGTGCAGCAACATGAA-GCAGTGTGTGGACTCCAATGCCTATGTGGCCT 2677
Db 2374 GAGTGGCATGTGGTGCAGCAACATGAAGGCAGTGTGTGGACTCCAATGCCTATGTGGCCT 2433
Qy 2678 CCTTCCCTTTT-GGCCAGTGTATGGAATGGTATACGATGA-GCACCTGCCCCCCTGAAAA 2735
Db 2434 CCTTCCCTTTTGGGCCAGTGTATGGAATGGTATACGATGAGGCACCTGCCCCCCTGAAAA 2493
Qy 2736 TTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCA-GGCTGTGGCTGGTGTA 2794
Db 2494 TTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGGCTGTGGCTGGTGTA 2553
Qy 2795 CTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGA 2854
Db 2554 CTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGA 2613
Qy 2855 AGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCCTGCTCAATTCCAGCA 2914
Db 2614 AGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCCTGCTCAATTCCAGCA 2673
Qy 2915 TGTGTCTAGAGGACAGCAGATACAACTGG--TCTTTCATTCACTGTCCAGCTTGCCAATG 2972
Db 2674 TGTGTCTAGAGGACAGCAGATACAACTGGGTCTTTCATTTCACTGTCCAGTTTGCCAATG 2733
Qy 2973 CAAC-GGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCA 3031
Db 2734 GAACGGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCA 2793
Qy 3032 CAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTAC-GGTGATCCCACCAA-TGGAGG 3089
Db 2794 CAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGGTGATCCCACCAATTGGAGG 2853
Qy 3090 GAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAA 3149
Db 2854 GAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAA 2913
Qy 3150 GTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAA 3209
Db 2914 GTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAA 2973
Qy 3210 TCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCA 3269
Db 2974 TCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCA 3033
Qy 3270 GTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGC 3329
Db 3034 GTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGC 3093
Qy 3330 TACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAA 3389
Db 3094 TACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAA 3153
Qy 3390 CCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCC 3449
Db 3154 CCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCC 3213
Qy 3450 TGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGA 3509
Db 3214 TGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGA 3273
Qy 3510 TTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCAT 3569
Db 3274 TTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCAT 3333

```
Qy       3570 CAAAATTCAGAT 3581
              |||||||||| |
Db       3334 CAAAATTCAGGT 3345
```

RESULT 13
AAZ91922
ID AAZ91922 standard; cDNA; 2625 BP.
XX
AC AAZ91922;
XX
DT 08-JUN-2000 (first entry)
XX
DE Human mahogany protein coding sequence #4.
XX
KW Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;
KW weight regulation; cell therapy; body weight disorder; cachexia;
KW anorexia; hyperpigmentation; increased metabolic rate disorder;
KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.
XX
OS Homo sapiens.
XX
PN WO200005373-A2.
XX
PD 03-FEB-2000.
XX
PF 21-JUL-1999; 99WO-US016484.
XX
PR 21-JUL-1998; 98US-0093630P.
PR 20-OCT-1998; 98US-0104978P.
PR 05-FEB-1999; 99US-00245041.
XX
PA (MILL-) MILLENIUM PHARM INC.
XX
PI Moore K, Nagle DL;
XX
DR WPI; 2000-195103/17.
DR P-PSDB; AAY81809.
XX
PT New human and murine mahogany genes, useful, e.g. for diagnosis and
PT treatment of body weight disorders.
XX
PS Claim 1; Fig 20a; 188pp; English.
XX
CC This sequence represents a human mahogany gene of the invention. The
CC mahogany genes are used: (i) to produce recombinant mahogany (mg)
CC proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming
CC therapeutics; (iii) as a source of diagnostic probes and primers for
CC detecting expression of mg genes or mutations, regulatory defects, in
CC this gene, or for isolation of related sequences; and (iv) in (cell-
CC based) gene therapy. (II) are used to raise specific antibodies (Ab); to
CC identify other (extra)cellular products involved in weight regulation,
CC and to screen for agents that disrupt interaction between (II) and other
CC macromolecules. The Ab are used to detect abnormal levels (or function)
CC of (II) (for diagnosis, prognosis or monitoring of treatment); to
CC evaluate (II)-expressing cells intended for cell therapy, and as
CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the
CC mg polypeptide) are used to identify agents (A) that modulate mg
CC activity. (A) are potentially useful for the treatment of body weight
CC disorders, particularly obesity, cachexia or anorexia, or other
CC conditions associated with the mg gene such as hyperpigmentation,
CC hyperphagia and disorders that result in increased metabolic rate
XX
SQ Sequence 2625 BP; 660 A; 605 C; 694 G; 666 T; 0 U; 0 Other;

```
  Query Match             50.4%;  Score 2051.8;  DB 3;  Length 2625;
  Best Local Similarity   97.1%;  Pred. No. 0;
  Matches 2089;  Conservative    0;  Mismatches   62;  Indels    0;  Gaps    0;

Qy         15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74
              || |||    |  ||   | ||  |  |  |  | | |   |||| ||   || || ||
Db        234 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAA 293

Qy         75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134
              |    | |  | ||  ||||||||||||||||||||||||||||||||||||||||||||
Db        294 GGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 353

Qy        135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        354 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 413

Qy        195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 473

Qy        255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 533

Qy        315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 593

Qy        375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 653

Qy        435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        654 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 713

Qy        495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        714 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 773

Qy        555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        774 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 833

Qy        615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        834 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 893

Qy        675 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        894 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 953

Qy        735 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        954 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 1013

Qy        795 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1014 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 1073

Qy        855 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1074 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 1133

Qy        915 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1134 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 1193

Qy        975 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1194 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1253

Qy       1035 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1254 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1313

Qy       1095 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1314 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1373

Qy       1155 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1214
```

```
Db       1374 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1433

Qy       1215 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1274

Db       1434 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1493

Qy       1275 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1334

Db       1494 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1553

Qy       1335 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1394

Db       1554 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1613

Qy       1395 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1454

Db       1614 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1673

Qy       1455 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1514

Db       1674 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1733

Qy       1515 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1574

Db       1734 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1793

Qy       1575 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGC 1634

Db       1794 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGC 1853

Qy       1635 AGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGA 1694

Db       1854 AGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGA 1913

Qy       1695 CATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGC 1754

Db       1914 CATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGC 1973

Qy       1755 AGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGC 1814

Db       1974 AGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGC 2033

Qy       1815 GCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCT 1874

Db       2034 GCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCT 2093

Qy       1875 TGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGA 1934

Db       2094 TGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGA 2153

Qy       1935 CTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGAT 1994

Db       2154 CTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGAT 2213

Qy       1995 CTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAA 2054

Db       2214 CTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAA 2273

Qy       2055 GACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGA 2114

Db       2274 GACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGA 2333

Qy       2115 GTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGG 2165

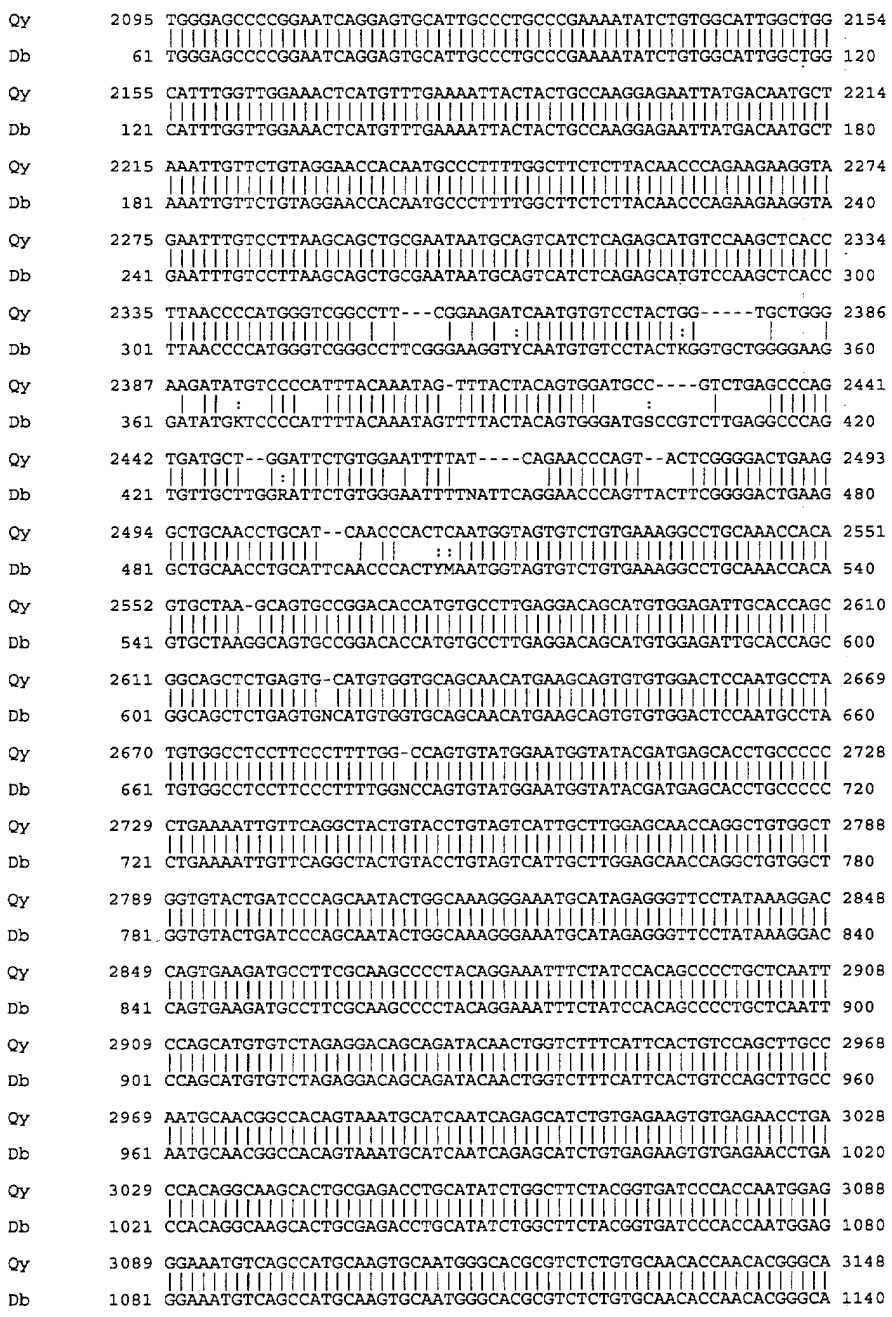
Db       2334 GTGCATTGCCCTGCCCGGTAGGCCTTGCAGGGTCATCTTGGTGTGTGTGGG 2384
```

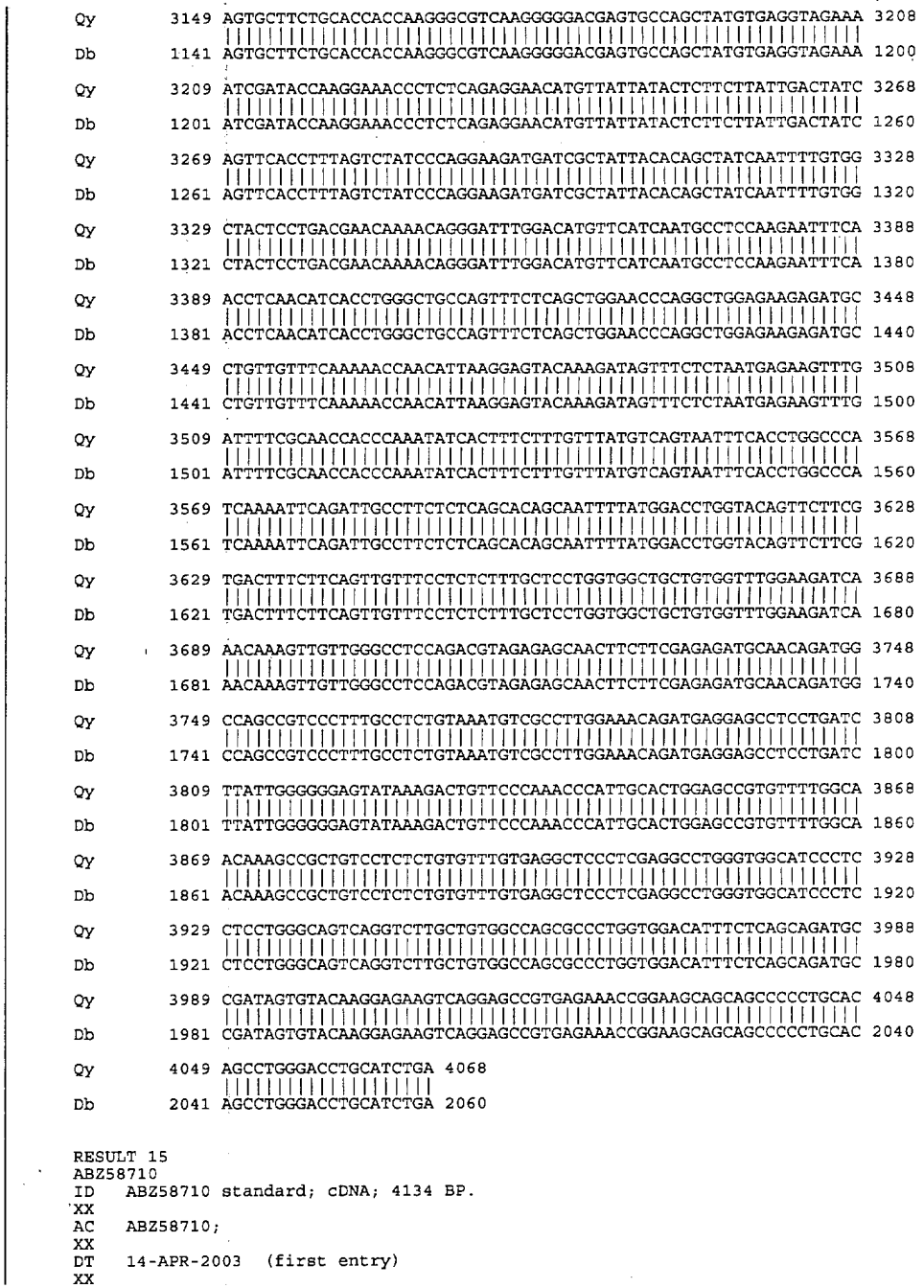
RESULT 14
AAZ91919
ID AAZ91919 standard; cDNA; 6373 BP.
XX
AC AAZ91919;
XX
DT 08-JUN-2000 (first entry)
XX
DE Human mahogany protein coding sequence #1.
XX
KW Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;
KW weight regulation; cell therapy; body weight disorder; cachexia;
KW anorexia; hyperpigmentation; increased metabolic rate disorder;
KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT misc_difference 448
FT /*tag= a
FT /note= "represented in specification as :"
FT misc_difference 615
FT /*tag= b
FT /note= "represented in specification as :"
FT misc_difference 683
FT /*tag= c
FT /note= "represented in specification as :"
XX
PN WO200005373-A2.
XX
PD 03-FEB-2000.
XX
PF 21-JUL-1999; 99WO-US016484.
XX
PR 21-JUL-1998; 98US-0093630P.
PR 20-OCT-1998; 98US-0104978P.
PR 05-FEB-1999; 99US-00245041.
XX
PA (MILL-) MILLENIUM PHARM INC.
XX
PI Moore K, Nagle DL;
XX
DR WPI; 2000-195103/17.
DR P-PSDB; AAY81806.
XX
PT New human and murine mahogany genes, useful, e.g. for diagnosis and
PT treatment of body weight disorders.
XX
PS Claim 1; Fig 10a; 188pp; English.
XX
CC This sequence represents a human mahogany gene of the invention. The
CC mahogany genes are used: (i) to produce recombinant mahogany (mg)
CC proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming
CC therapeutics; (iii) as a source of diagnostic probes and primers for
CC detecting expression of mg genes or mutations, regulatory defects, in
CC this gene, or for isolation of related sequences; and (iv) in (cell-
CC based) gene therapy. (II) are used to raise specific antibodies (Ab); to
CC identify other (extra)cellular products involved in weight regulation,
CC and to screen for agents that disrupt interaction between (II) and other
CC macromolecules. The Ab are used to detect abnormal levels (or function)
CC of (II) (for diagnosis, prognosis or monitoring of treatment); to
CC evaluate (II)-expressing cells intended for cell therapy, and as
CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the
CC mg polypeptide) are used to identify agents (A) that modulate mg
CC activity. (A) are potentially useful for the treatment of body weight
CC disorders, particularly obesity, cachexia or anorexia, or other
CC conditions associated with the mg gene such as hyperpigmentation,
CC hyperphagia and disorders that result in increased metabolic rate
XX
SQ Sequence 6373 BP; 1736 A; 1468 C; 1460 G; 1699 T; 0 U; 10 Other;

Query Match 45.0%; Score 1830.2; DB 3; Length 6373;
Best Local Similarity 96.4%; Pred. No. 0;
Matches 1986; Conservative 7; Mismatches 41; Indels 26; Gaps 11;

```
Qy       2035 ATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAG 2094

Db          1 ATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAG 60
```





RESULT 15
ABZ58710
ID ABZ58710 standard; cDNA; 4134 BP.
XX
AC ABZ58710;
XX
DT 14-APR-2003 (first entry)
XX

```
DE   Murine homologue of attractin/mahogany (HAM) polypeptide encoding cDNA.
XX
KW   HAM; homologue of attractin/mahogany; immunosuppressive; cytostatic;
KW   antiinflammatory; cardiant; osteopathic; gene therapy; mouse; gene; ss.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .4134
FT                   /*tag=  a
FT                   /product= "HAM"
FT                   /note= "homologue of attractin/mahogany"
XX
PN   WO200297120-A1.
XX
PD   05-DEC-2002.
XX
PF   23-MAY-2002; 2002WO-US016391.
XX
PR   25-MAY-2001; 2001US-0293608P.
PR   24-SEP-2001; 2001US-0324626P.
XX
PA   (IMMV ) IMMUNEX CORP.
XX
PI   Anderson DM;
XX
DR   WPI; 2003-140486/13.
DR   P-PSDB; ABP71102.
XX
PT   New Homologue of Attractin/Mahagony (HAM) polypeptide, useful for
PT   treating HAM-associated disorder consisting of inflammatory, autoimmune,
PT   cell proliferative or cardiovascular disorders.
XX
PS   Claim 14; Fig 5A-D; 89pp; English.
XX
CC   The invention relates to Homologue of Attractin/Mahogany (HAM)
CC   polypeptides and encoding polynucleotides. The HAM polypeptides can be
CC   expressed by standard recombinant methodology. The HAM polypeptides are
CC   useful for treating HAM-associated disorder consisting of inflammatory,
CC   autoimmune, graft-versus-host, neurological, myelination, cell
CC   proliferative, cardiovascular, haematologic, liver, metabolic, weight or
CC   bone disorder. The present sequence represents a cDNA encoding a mouse
CC   HAM polypeptide
XX
SQ   Sequence 4134 BP; 1085 A; 981 C; 1036 G; 1032 T; 0 U; 0 Other;

  Query Match             36.4%;  Score 1481;  DB 7;  Length 4134;
  Best Local Similarity   63.4%;  Pred. No. 0;
  Matches 2507;  Conservative    0;  Mismatches 1360;  Indels   90;  Gaps   12;

Qy        117 TGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGG 176
              || || |  |  ||| | |||  | | |||||||| || || ||||||||||||
Db        219 TGTCAACTCCACCTGCCTGTGCGACCCGGGCTGGGTTGGGGACCAGTGCCAGCACTGCCA 278

Qy        177 GGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAA 236
              ||||  | ||||   |||| | | |||||||  | | | ||||||||||||    || |||||
Db        279 GGGCAGGTTCAAGTTAACAGAACCTTCTGGATATTTAACAGATGGACCAATTAACTATAA 338

Qy        237 ATACAAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCG 296
              ||| ||||| ||||| || ||||| ||||||||  | ||||||  | |  | ||  |  |
Db        339 ATATAAAACAAAGTGTACATGGCTAATTGAAGGCTATCCAAATGCAGTGCTAAGGTTAAG 398

Qy        297 TTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTC 356
              |||||||||||||||||||| || || ||||| ||| | ||||||||||||||| || ||
Db        399 ATTCAATCATTTTGCTACAGAATGCAGCTGGGATCATATGTATGTTTATGATGGAGATTC 458

Qy        357 AATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAA 416
              || || |||||  |||| |||| | |||||||| | || ||||||||        || ||
Db        459 TATATACGCACCTTTAGTAGCTGTACTTAGTGGCTTGATCGTTCCTGAAGTGAGGGGTAA 518

Qy        417 TGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGA 476
              ||||| || |||||||| ||  | || || || || ||  |||| || ||||| || ||
Db        519 CGAGACCGTGCCTGAGGTGGTCACGACGTCTGGCTACGCGCTGCTCCACTTTTTCAGCGA 578

Qy        477 TGCTGCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAA 536
              |||||| ||||| | |||| || || ||| |||  |||       |  ||  ||||| || ||
Db        579 TGCTGCATATAACCTAACTGGCTTCAACATTTTTTATTCGATCAATTCCTGTCCTAACAA 638

Qy        537 CTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACT---GTTGAATGTGA 593
              |||||| || |  ||| ||||||  | ||||          | |  |   ||  | |||||
Db        639 CTGCTCTGGTCATGGAAAGTGTACAACCAGTGTCTCTGTTGCAAGTCAAGTGTATTGTGA 698

Qy        594 ATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGG 653
              |||     || |||||||||| |||||||||||||||||||  |||||  ||  ||| |||||
Db        699 ATGCGACAAATACTGGAAAGGGGAAGCATGTGACATTCCTTACTGTAAAGCCAATTGTGG 758

Qy        654 TTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGA 713
              |||  |||  |||  |||  |  ||  |  |  |  |  ||  |||| |  |
Db        759 GAGTCCAGATCATGGCTACTGTGACCTAACAGGAGAGAAACTCTGTGTCTGCAACGATAG 818

Qy        714 CTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGA 773
              |||||  ||  ||  |  |||||  |   |||  ||  |||   ||||  |  |||| |
Db        819 TTGGCAAGGCCCAGATTGTTCTCTGAATGTCCCTTCTACTGAGTCTTACTGGATTTTGCC 878

Qy        774 GGAATATTCTAACTTAAAGCTCCCC------AGAGCATCTCATAAAGCTGTGGTCAATGG 827
              |   |      |||  |  |  ||        || || ||||||||  ||  |  ||||
Db        879 AAATGTTAAACCCTTCAGCCCTTCCGTAGGTCGGGCCTCACATAAAGCAGTTTTACATGG 938

Qy        828 AAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCT 887
              ||  | ||||||||  ||||||||||||  ||  ||| ||    ||  |  ||||||||
Db        939 GAAATTCATGTGGGTGATTGGAGGATATACTTTTAACTACAGTTCTTTTCAAATGGTTCT 998

Qy        888 AGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGT 947
              |  ||  | | |  ||  |||  |  |  ||  |  |  |  |  |  |  |
Db        999 GAATTACAATTTAGAAAGCAGTATATGGAATGTAGGTGCTGTATCAAGGGGCCCT---CT 1055

Qy        948 TGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAA 1007
              |  |||||  ||  ||||||  |  ||  |  ||  ||||  ||  ||  |  |||||  |||||  |
Db       1056 TCAGAGATACGGACATTCTCTTGCCCTCTATCAGGAAAACATCTTTATGTACGGAGGCAG 1115

Qy       1008 AATTGATTCAA---CTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGA 1064
              |||  ||  |  |  |||  ||  ||  |||  ||||  ||  |  ||  |||  ||  |  |  |  |
Db       1116 AATGGAGACCAGTGATGGCAACGTCACCGATGAATTATGGGTGTTTAACGTACGCAGTCA 1175

Qy       1065 GTCATGGGTGTTGTTGACCCCTAAGGCAAAGG---------AGCAGTATGCAGTGGTTGG 1115
              ||||||  |  |||||  |  |  |  |  |||||||  ||  ||||  ||
Db       1176 ATCATGGAGCACGAAAACCCCCACTGTCCTTGGCCACAGTCAGCAGTACGCTGTGGAGGG 1235

Qy       1116 GCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGG 1175
              |||||  ||||||||  |  |||  |  |  |  |  |||||||||  |  |||||  |||||
Db       1236 ACACTCGGCACACATCATGGAGCTGGACAGTAGAGACGTGGTCATGATTGTCATATTTGG 1295

Qy       1176 TCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACAC 1235
              |  |  |  |  |||||  |||  ||||  | ||||||||  ||  |  |||||
Db       1296 ATATTCTGCAATATATGGCTATACCAGCAGCATACAGGAATACCATATCTCCTCAAACAC 1355

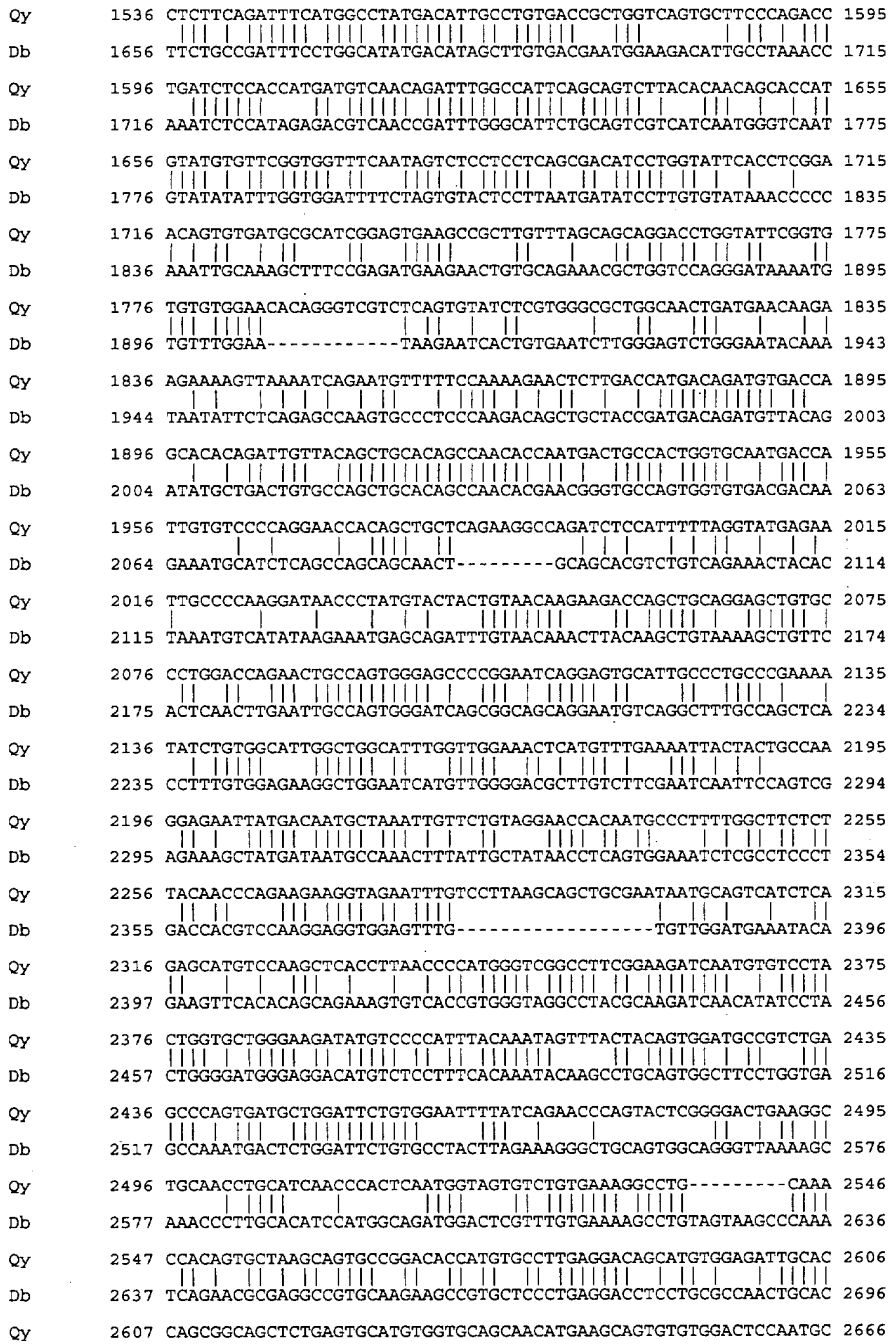
Qy       1236 ATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGT 1295
              |||  |  |  |  |||  |||||||||||  ||  ||||  |  |||||  ||||
Db       1356 TTGGCTAGTTCCAGAAACGAAAGGAGCCATTGTGCAAGGTGGATATGGCCACACAAGTGT 1415

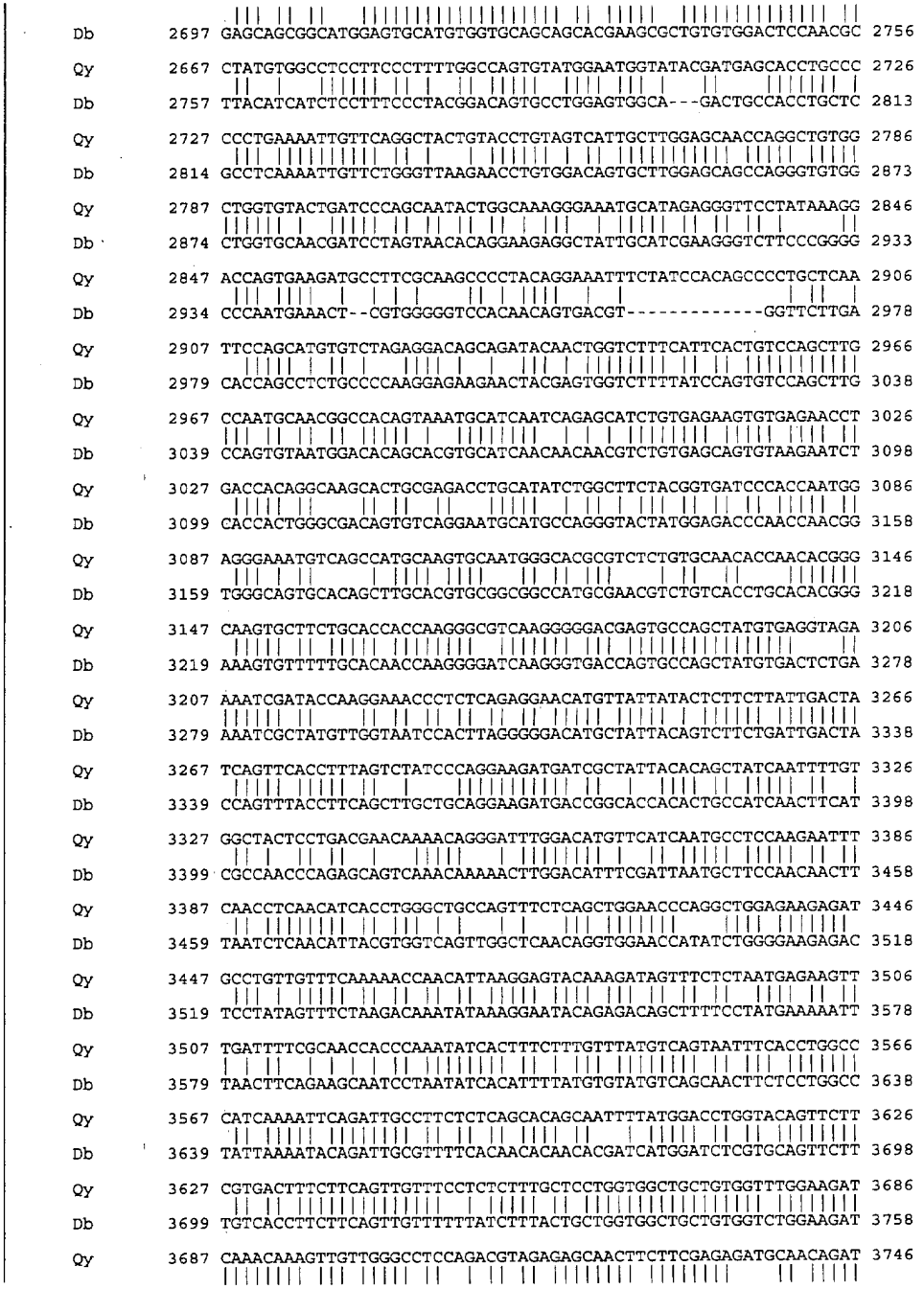
Qy       1296 TTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAA 1355
              ||  |||  |  |||||  |  ||  |  ||||||  ||  ||||||||  ||  ||  ||  |||
Db       1416 GTATGACGAAGTGACCAAGTCCATCTATGTTCACGGAGGCTACAAAGCATTGCCTGGCAA 1475

Qy       1356 TAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCAT 1415
              |||||||  ||||  |  ||  ||  |||||  ||  ||  ||  |  |||  ||  |||||  ||
Db       1476 TAAGTACGGGCTGGTGGACGACCTCTATAAGTACGAAGTCAACACCAGGACTTGGACTAT 1535

Qy       1416 TCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACCAT 1475
              |  |  ||  ||  |  |||  |  |||  |  |||  |||||||  |  |  ||||  |  ||
Db       1536 TTTGAAAGAAAGTGGGTTTGCCAGATACCTTCACTCAGCTGTTCTTATCAATGGAGCTAT 1595

Qy       1476 GCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATGCTT 1535
              |||  |  |||||  |||||  ||  ||  |||||  ||  ||  ||||  |  ||  ||  |||||  ||
Db       1596 GCTTATTTTTGGAGGAAATACCCATAATGATACTTCCCTGAGTAACGGTGCAAAATGTTT 1655
```





```
Db       3759 CAAACAAACTTGCTGGGCTTCTCGTCGGAGGGAGCAACTGCTTCGAGAACGACAGCAGAT 3818

Qy       3747 GGCCAGCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGAGGAGCCTCCTGA 3806
             |||||||||||||||||| |||||  |||| ||| |||||  ||  |  || |   | ||
Db       3819 GGCCAGCCGTCCCTTTGCTTCTGTTGATGTAGCCCTGGAAGTAGGAGCTGAACAGACAGA 3878

Qy       3807 TCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGAGCCGTGTTTTGG 3866
               || |  | |||    || ||  ||  || || || || ||  | || || ||   |||
Db       3879 CTTTCTGCGAGGGCCATTAGAGGGTGCCCCTAAGCCAATAGCCATCGAACCCTGCGCTGG 3938

Qy       3867 CAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCTGGGTGGCATCCC 3926
              |||| ||| |||||||  |||||||| |  | || ||  |||        |||  |||
Db       3939 GAACAGAGCTGCTGTCCTGACTGTGTTTCTCTGTCTACCGAGAGGATCTTCAGGCGCCCC 3998

Qy       3927 TCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGAT 3986
              || |||||||||||||| |||||| | ||||| |||||| | ||||| || |||||||
Db       3999 ACCCCCTGGGCAGTCAGGCCTTGCTATCGCCAGTGCCCTGATAGACATCTCACAGCAGAA 4058

Qy       3987 GCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCAGCAGCCCCC 4043
              |||    |   | || || ||| |    |  ||  | || |||||||| |   || |
Db       4059 GCCTTCTGATAATAAAGACAAGACTTCTGGAGTCCGCAATCGGAAGCACCTCTCCAC 4115
```

Search completed: March 4, 2004, 18:09:10
Job time : 971.587 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2004, 17:06:39 ; Search time 899.014 Seconds
(without alignments)
16549.723 Million cell updates/sec

Title: US-09-787-097-11
Perfect score: 4068
Sequence: 1 atggtggccgcagcggcggc.....agcctgggacctgcatctga 4068

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2421054 seqs, 1828716029 residues

Total number of hits satisfying chosen parameters: 4842108

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	3988.4	98.0	8589	9	US-09-893-238-14	Sequence 14, Appl
	2	3499.8	86.0	4072	9	US-09-893-238-16	Sequence 16, Appl
	3	3326.4	81.8	8827	9	US-09-893-238-1	Sequence 1, Appli
	4	2051.8	50.4	2625	9	US-09-893-238-18	Sequence 18, Appl
	5	1853	45.6	6370	9	US-09-893-238-12	Sequence 12, Appl
	6	1461.8	35.9	6733	16	US-10-197-824-1	Sequence 1, Appli
	7	1451.8	35.7	2419	9	US-09-893-238-8	Sequence 8, Appli
	8	463	11.4	1051	9	US-09-893-238-10	Sequence 10, Appl
c	9	277.4	6.8	625	14	US-10-198-846-5869	Sequence 5869, Ap
	10	250.8	6.2	531	9	US-09-864-761-15900	Sequence 15900, A
c	11	246	6.0	246	9	US-09-736-457-966	Sequence 966, App
c	12	246	6.0	246	9	US-09-902-941-966	Sequence 966, App
c	13	246	6.0	246	9	US-09-849-626-966	Sequence 966, App
c	14	246	6.0	246	14	US-10-017-754-966	Sequence 966, App
c	15	246	6.0	246	14	US-10-113-872-966	Sequence 966, App
	16	234.4	5.8	207433	15	US-10-277-216-5	Sequence 5, Appli
	17	234.4	5.8	207433	16	US-10-126-022-5	Sequence 5, Appli
	18	212.8	5.2	415	9	US-09-833-381-1916	Sequence 1916, Ap
	19	202.8	5.0	90050	9	US-09-893-238-5	Sequence 5, Appli
	20	201.2	4.9	212	9	US-09-796-692-7610	Sequence 7610, Ap
	21	201.2	4.9	212	14	US-10-040-862-7610	Sequence 7610, Ap
	22	201.2	4.9	212	15	US-10-057-475B-7610	Sequence 7610, Ap
	23	201.2	4.9	212	15	US-10-154-884B-7610	Sequence 7610, Ap
	24	193	4.7	193	9	US-09-864-761-32404	Sequence 32404, A
c	25	182	4.5	182	9	US-09-864-761-21460	Sequence 21460, A
c	26	174.6	4.3	375	9	US-09-864-761-4720	Sequence 4720, Ap
c	27	169.2	4.2	17056	9	US-09-893-238-3	Sequence 3, Appli
c	28	141.4	3.5	402	9	US-09-864-761-3681	Sequence 3681, Ap
c	29	137	3.4	137	9	US-09-864-761-20449	Sequence 20449, A
c	30	132.2	3.2	226	9	US-09-893-238-6	Sequence 6, Appli
c	31	115.4	2.8	221	9	US-09-893-238-7	Sequence 7, Appli
	32	100.2	2.5	5973	9	US-09-893-238-4	Sequence 4, Appli
	33	89	2.2	594	15	US-10-027-632-264639	Sequence 264639,
	34	89	2.2	594	15	US-10-027-632-264640	Sequence 264640,
	35	88.4	2.2	629	15	US-10-027-632-270240	Sequence 270240,
	36	71.8	1.8	713	13	US-10-040-739-127	Sequence 127, App
	37	71.6	1.8	5312	14	US-10-311-455-169	Sequence 169, App
	38	60	1.5	60	10	US-09-908-975-15421	Sequence 15421, A
	39	58.6	1.4	65	10	US-09-908-975-29990	Sequence 29990, A
c	40	54.6	1.3	468	15	US-10-027-632-108988	Sequence 108988,
	41	48.8	1.2	1731	12	US-10-282-122A-28363	Sequence 28363, A
	42	46.2	1.1	2004	12	US-10-282-122A-28310	Sequence 28310, A
c	43	46	1.1	2561	9	US-09-976-740-48	Sequence 48, Appl
c	44	46	1.1	2561	12	US-10-671-242-48	Sequence 48, Appl
c	45	46	1.1	2561	13	US-10-023-529-48	Sequence 48, Appl

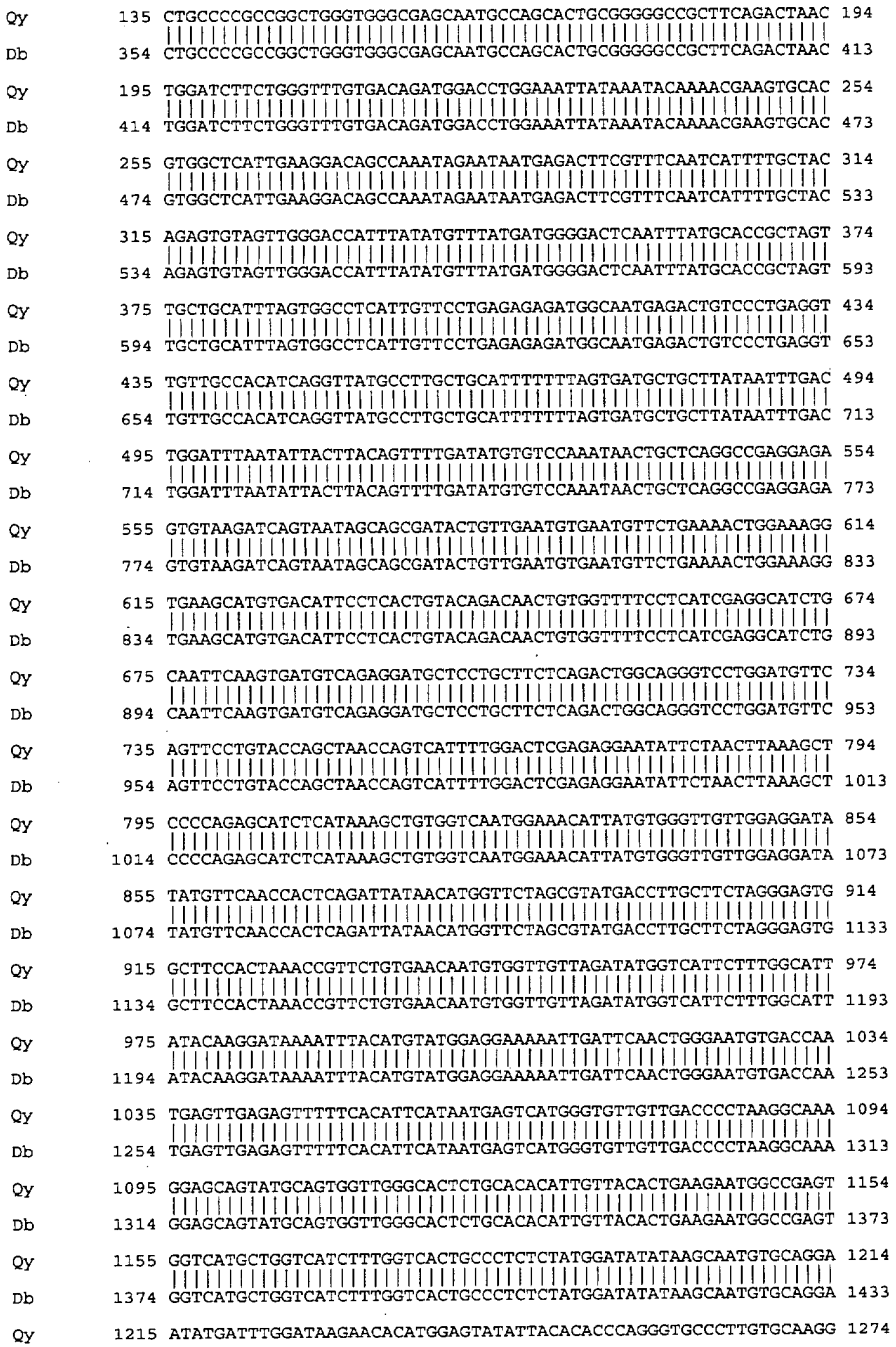
ALIGNMENTS

RESULT 1
US-09-893-238-14
; Sequence 14, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 8589
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-893-238-14

Query Match 98.0%; Score 3988.4; DB 9; Length 8589;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 4013; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```
Qy          15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74
               || |||     |  ||    | || |  |  |  | | |    ||||  ||   || || ||
Db         234 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAA 293

Qy          75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134
               |    | |  | ||  ||||||||||||||||||||||||||||||||||||||||||||
Db         294 GGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 353
```


Qy 135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194
Db 354 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 413
Qy 195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254
Db 414 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 473
Qy 255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314
Db 474 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 533
Qy 315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374
Db 534 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 593
Qy 375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434
Db 594 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 653
Qy 435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494
Db 654 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 713
Qy 495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554
Db 714 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 773
Qy 555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614
Db 774 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 833
Qy 615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 674
Db 834 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 893
Qy 675 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 734
Db 894 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 953
Qy 735 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 794
Db 954 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 1013
Qy 795 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 854
Db 1014 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 1073
Qy 855 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 914
Db 1074 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 1133
Qy 915 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 974
Db 1134 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 1193
Qy 975 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1034
Db 1194 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1253
Qy 1035 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1094
Db 1254 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1313
Qy 1095 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1154
Db 1314 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1373
Qy 1155 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1214
Db 1374 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1433
Qy 1215 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1274
Db 1434 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1493
Qy 1275 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1334
Db 1494 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1553
Qy 1335 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1394
Db 1554 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1613
Qy 1395 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1454
Db 1614 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1673
Qy 1455 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1514
Db 1674 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1733
Qy 1515 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1574
Db 1734 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1793
Qy 1575 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGC 1634
Db 1794 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGC 1853
Qy 1635 AGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGA 1694
Db 1854 AGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGA 1913
Qy 1695 CATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGC 1754
Db 1914 CATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGC 1973
Qy 1755 AGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGC 1814
Db 1974 AGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGC 2033
Qy 1815 GCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCT 1874
Db 2034 GCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCT 2093
Qy 1875 TGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGA 1934
Db 2094 TGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGA 2153
Qy 1935 CTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGAT 1994
Db 2154 CTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGAT 2213
Qy 1995 CTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAA 2054
Db 2214 CTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAA 2273
Qy 2055 GACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGA 2114
Db 2274 GACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGA 2333
Qy 2115 GTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATG 2174
Db 2334 GTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATG 2393
Qy 2175 TTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCA 2234
Db 2394 TTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCA 2453
Qy 2235 CAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCT 2294
Db 2454 CAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCT 2513
Qy 2295 GCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCT 2354

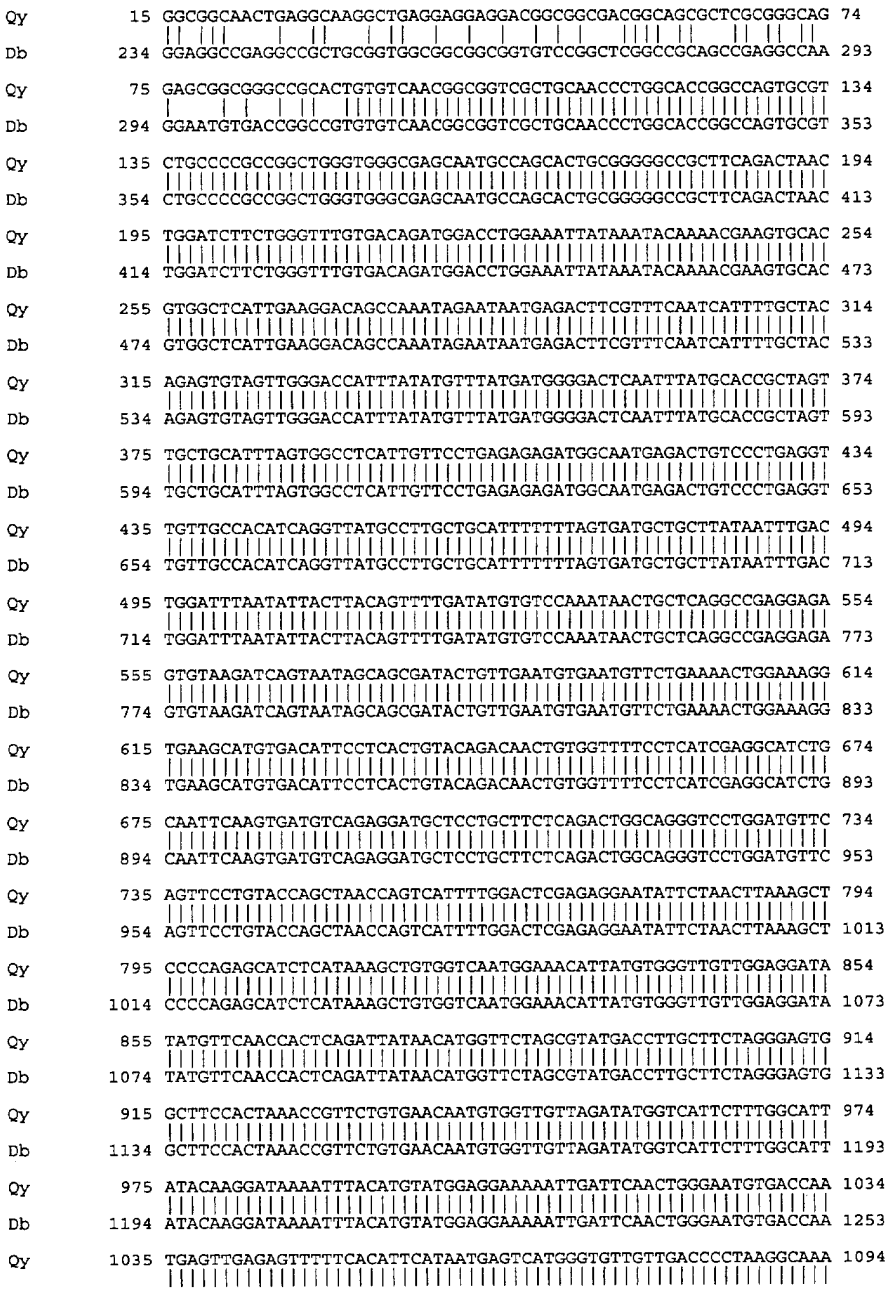
```
Db       2514 GCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCT 2573
Qy       2355 TCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTT 2414
Db       2574 TCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTT 2633
Qy       2415 ACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACC 2474
Db       2634 ACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACC 2693
Qy       2475 CAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGA 2534
Db       2694 CAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGA 2753
Qy       2535 AAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATG 2594
Db       2754 AAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATG 2813
Qy       2595 TGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGT 2654
Db       2814 TGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGT 2873
Qy       2655 GGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGAT 2714
Db       2874 GGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGAT 2933
Qy       2715 GAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCA 2774
Db       2934 GAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCA 2993
Qy       2775 ACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGG 2834
Db       2994 ACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGG 3053
Qy       2835 TTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACA 2894
Db       3054 TTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACA 3113
Qy       2895 GCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCA 2954
Db       3114 GCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCA 3173
Qy       2955 CTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAA 3014
Db       3174 CTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAA 3233
Qy       3015 GTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGA 3074
Db       3234 GTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGA 3293
Qy       3075 TCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAA 3134
Db       3294 TCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAA 3353
Qy       3135 CACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCT 3194
Db       3354 CACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCT 3413
Qy       3195 ATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCT 3254
Db       3414 ATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCT 3473
Qy       3255 TCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGC 3314
Db       3474 TCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGC 3533
Qy       3315 TATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGC 3374
Db       3534 TATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGC 3593
Qy       3375 CTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGC 3434
Db       3594 CTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGC 3653

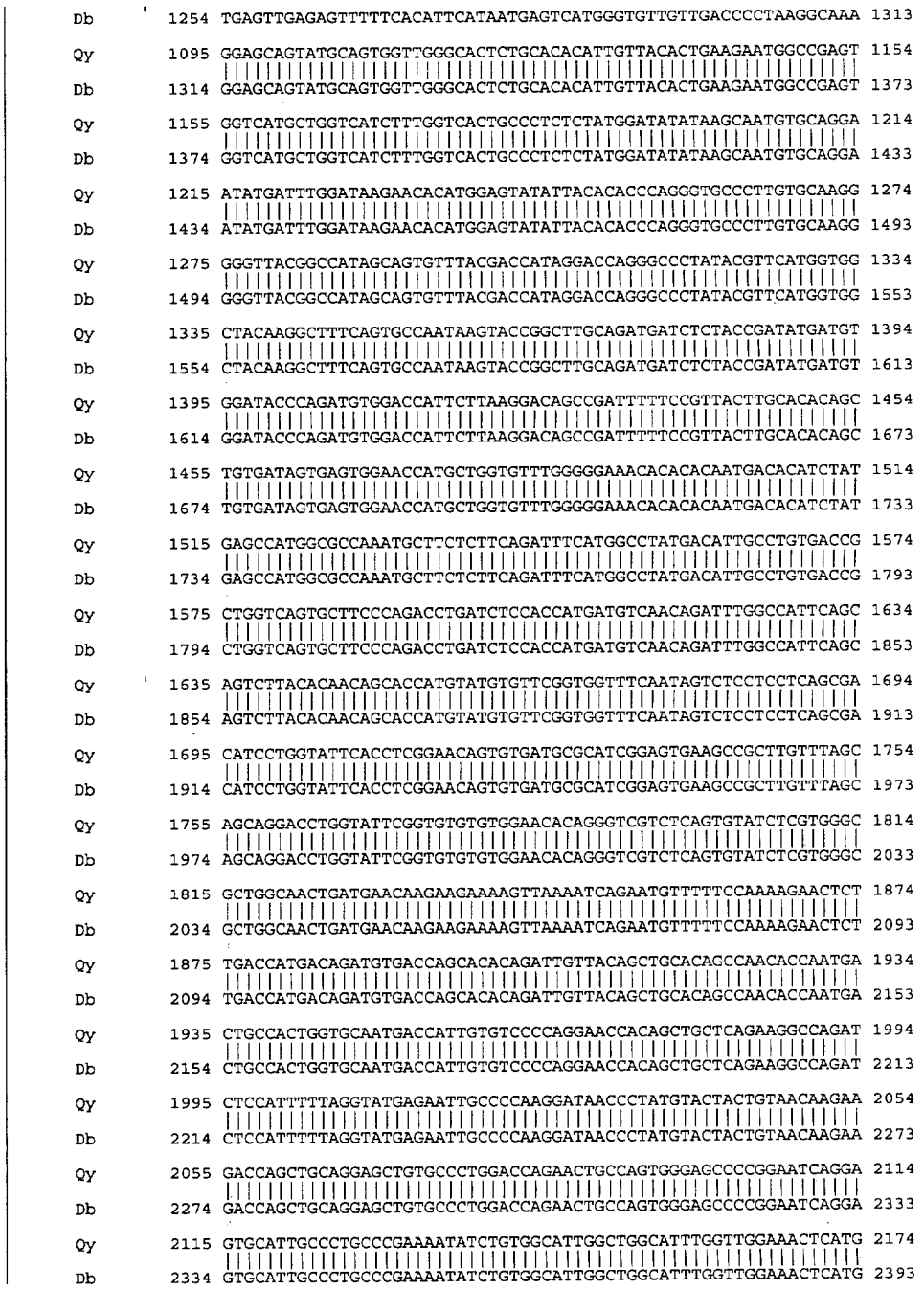
Qy       3435 TGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTC 3494
Db       3654 TGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTC 3713
Qy       3495 TAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAA 3554
Db       3714 TAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAA 3773
Qy       3555 TTTCACCTGGCCCATCAAAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCT 3614
Db       3774 TTTCACCTGGCCCATCAAAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCT 3833
Qy       3615 GGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGT 3674
Db       3834 GGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGT 3893
Qy       3675 GGTTTGGAAGATCAAACAAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGA 3734
Db       3894 GGTTTGGAAGATCAAACAAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGA 3953
Qy       3735 GATGCAACAGATGGCCAGCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGA 3794
Db       3954 GATGCAACAGATGGCCAGCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGA 4013
Qy       3795 GGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGA 3854
Db       4014 GGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGA 4073
Qy       3855 GCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCT 3914
Db       4074 GCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCT 4133
Qy       3915 GGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACAT 3974
Db       4134 GGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACAT 4193
Qy       3975 TTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCA 4034
Db       4194 TTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCA 4253
Qy       4035 GCAGCCCCCTGCACAGCCTGGGACCTGCATCTGA 4068
Db       4254 GCAGCCCCCTGCACAGCCTGGGACCTGCATCTGA 4287
```

RESULT 2
US-09-893-238-16
; Sequence 16, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 4072
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-893-238-16

Query Match 86.0%; Score 3499.8; DB 9; Length 4072;

Best Local Similarity 98.8%; Pred. No. 0;
Matches 3525; Conservative 0; Mismatches 42; Indels 0; Gaps 0;





```
Qy      2175 TTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCA 2234
Db      2394 TTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCA 2453

Qy      2235 CAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCT 2294
Db      2454 CAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCT 2513

Qy      2295 GCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCT 2354
Db      2514 GCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCT 2573

Qy      2355 TCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTT 2414
Db      2574 TCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTT 2633

Qy      2415 ACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACC 2474
Db      2634 ACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACC 2693

Qy      2475 CAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGA 2534
Db      2694 CAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGA 2753

Qy      2535 AAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATG 2594
Db      2754 AAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATG 2813

Qy      2595 TGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGT 2654
Db      2814 TGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGT 2873

Qy      2655 GGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGAT 2714
Db      2874 GGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGAT 2933

Qy      2715 GAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCA 2774
Db      2934 GAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCA 2993

Qy      2775 ACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGG 2834
Db      2994 ACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGG 3053

Qy      2835 TTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACA 2894
Db      3054 TTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACA 3113

Qy      2895 GCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCA 2954
Db      3114 GCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCA 3173

Qy      2955 CTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAA 3014
Db      3174 CTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAA 3233

Qy      3015 GTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGA 3074
Db      3234 GTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGA 3293

Qy      3075 TCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAA 3134
Db      3294 TCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAA 3353

Qy      3135 CACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCT 3194
Db      3354 CACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCT 3413

Qy      3195 ATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCT 3254
Db      3414 ATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCT 3473

Qy      3255 TCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGC 3314
Db      3474 TCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGC 3533

Qy      3315 TATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGC 3374
Db      3534 TATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGC 3593

Qy      3375 CTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGC 3434
Db      3594 CTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGC 3653

Qy      3435 TGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTC 3494
Db      3654 TGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTC 3713

Qy      3495 TAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAA 3554
Db      3714 TAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAA 3773

Qy      3555 TTTCACCTGGCCCATCAAAATTCAGAT 3581
Db      3774 TTTCACCTGGCCCATCAAAATTCAGGT 3800
```

RESULT 3
US-09-893-238-1
; Sequence 1, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 8827
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-893-238-1

Query Match 81.8%; Score 3326.4; DB 9; Length 8827;
Best Local Similarity 91.1%; Pred. No. 0;
Matches 3534; Conservative 0; Mismatches 346; Indels 0; Gaps 0;

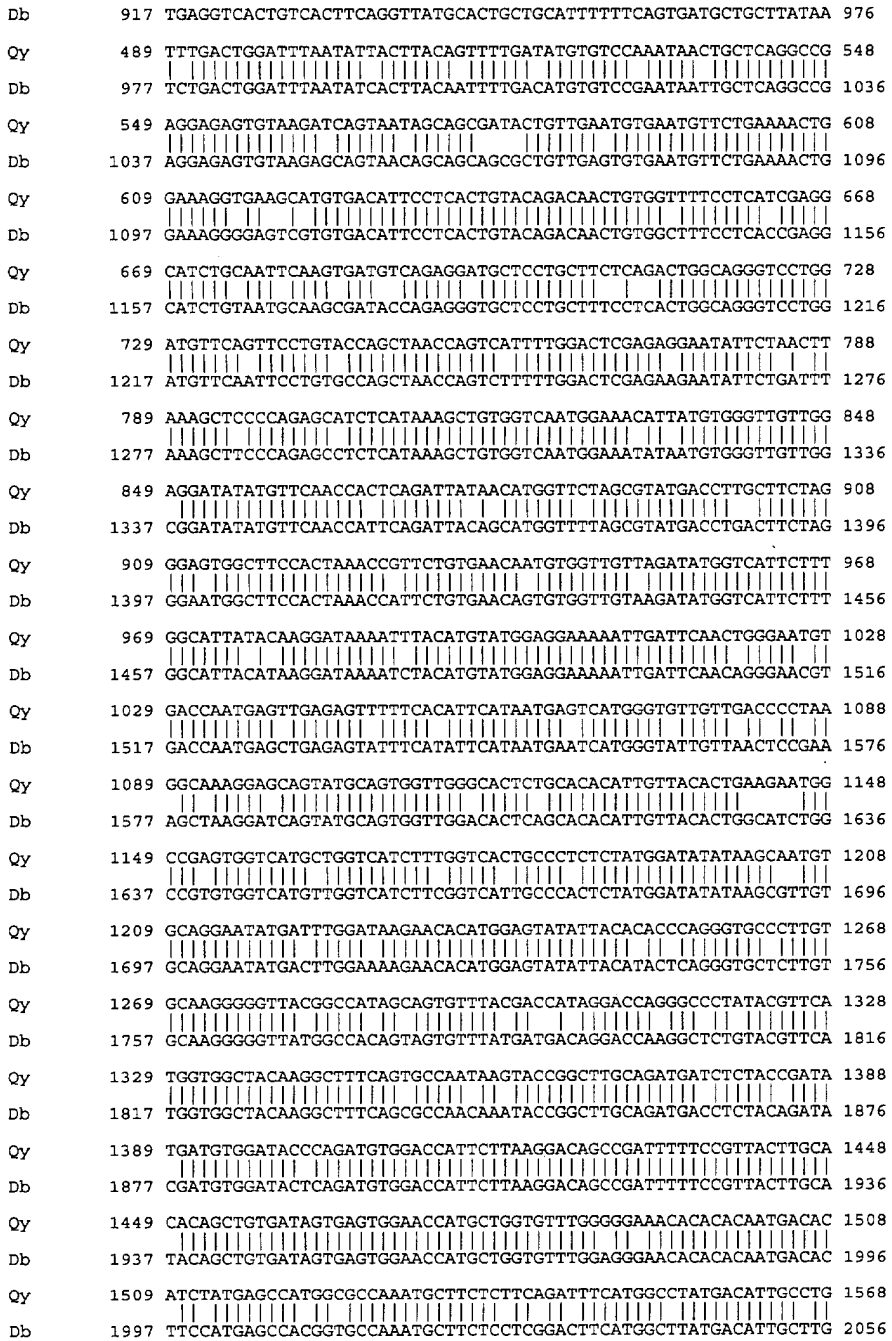
```
Qy       189 ACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAA 248
Db       677 ACTAACTGGCTCTTCTGGATTTGTAACAGATGGACCTGGGAATTATAAATATAAGACGAA 736

Qy       249 GTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTT 308
Db       737 GTGCACATGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGCTTCAACCATTT 796

Qy       309 TGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACC 368
Db       797 TGCTACAGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATCTACGCACC 856

Qy       369 GCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCC 428
Db       857 TCTGATTGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCTCC 916

Qy       429 TGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAA 488
```


Db 917 TGAGGTCACTGTCACTTCAGGTTATGCACTGCTGCATTTTTTCAGTGATGCTGCTTATAA 976
Qy 489 TTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCG 548
Db 977 TCTGACTGGATTTAATATCACTTACAATTTTGACATGTGTCCGAATAATTGCTCAGGCCG 1036
Qy 549 AGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTG 608
Db 1037 AGGAGAGTGTAAGAGCAGTAACAGCAGCAGCGCTGTTGAGTGTGAATGTTCTGAAAACTG 1096
Qy 609 GAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGG 668
Db 1097 GAAAGGGGAGTCGTGTGACATTCCTCACTGTACAGACAACTGTGGCTTTCCTCACCGAGG 1156
Qy 669 CATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGG 728
Db 1157 CATCTGTAATGCAAGCGATACCAGAGGGTGCTCCTGCTTTCCTCACTGGCAGGGTCCTGG 1216
Qy 729 ATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTT 788
Db 1217 ATGTTCAATTCCTGTGCCAGCTAACCAGTCTTTTTGGACTCGAGAAGAATATTCTGATTT 1276
Qy 789 AAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGG 848
Db 1277 AAAGCTTCCCAGAGCCTCTCATAAAGCTGTGGTCAATGGAAATATAATGTGGGTTGTTGG 1336
Qy 849 AGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAG 908
Db 1337 CGGATATATGTTCAACCATTCAGATTACAGCATGGTTTTAGCGTATGACCTGACTTCTAG 1396
Qy 909 GGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTT 968
Db 1397 GGAATGGCTTCCACTAAACCATTCTGTGAACAGTGTGGTTGTAAGATATGGTCATTCTTT 1456
Qy 969 GGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGT 1028
Db 1457 GGCATTACATAAGGATAAAATCTACATGTATGGAGGAAAAATTGATTCAACAGGGAACGT 1516
Qy 1029 GACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAA 1088
Db 1517 GACCAATGAGCTGAGAGTATTTCATATTCATAATGAATCATGGGTATTGTTAACTCCGAA 1576
Qy 1089 GGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGG 1148
Db 1577 AGCTAAGGATCAGTATGCAGTGGTTGGACACTCAGCACACATTGTTACACTGGCATCTGG 1636
Qy 1149 CCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGT 1208
Db 1637 CCGTGTGGTCATGTTGGTCATCTTCGGTCATTGCCCACTCTATGGATATATAAGCGTTGT 1696
Qy 1209 GCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGT 1268
Db 1697 GCAGGAATATGACTTGGAAAAGAACACATGGAGTATATTACATACTCAGGGTGCTCTTGT 1756
Qy 1269 GCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCA 1328
Db 1757 GCAAGGGGGTTATGGCCACAGTAGTGTTTATGATGACAGGACCAAGGCTCTGTACGTTCA 1816
Qy 1329 TGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATA 1388
Db 1817 TGGTGGCTACAAGGCTTTCAGCGCCAACAAATACCGGCTTGCAGATGACCTCTACAGATA 1876
Qy 1389 TGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCA 1448
Db 1877 CGATGTGGATACTCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCA 1936
Qy 1449 CACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACAC 1508
Db 1937 TACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGAGGGAACACACACAATGACAC 1996
Qy 1509 ATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTG 1568
Db 1997 TTCCATGAGCCACGGTGCCAAATGCTTCTCCTCGGACTTCATGGCTTATGACATTGCTTG 2056
Qy 1569 TGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCA 1628
Db 2057 TGACCGATGGTCAGTGCTTCCCAGACCTGAGCTCCATCATGATGTCAACAGATTTGGCCA 2116
Qy 1629 TTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCT 1688
Db 2117 TTCAGCAGTCTTGTACAACAGCACCATGTATGTGTTCGGCGGCTTCAACAGCCTCCTCCT 2176
Qy 1689 CAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTG 1748
Db 2177 CAGTGACGTCTTGGTCTTTACCTCGGAGCAGTGCGATGCACACCGCAGTGAAGCTGCTTG 2236
Qy 1749 TTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTC 1808
Db 2237 TGTGGCAGCAGGACCTGGTATCCGGTGTCTGTGGGACACACAGTCGTCTCGATGTACCTC 2296
Qy 1809 GTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAG 1868
Db 2297 CTGGGAGTTGGCAACTGAAGAACAAGCAGAAAAGTTAAAATCAGAGTGTTTTTCTAAAAG 2356
Qy 1869 AACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACAC 1928
Db 2357 AACCCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAATAC 2416
Qy 1929 CAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGG 1988
Db 2417 CAATGACTGCCACTGGTGCAATGATCACTGTGTCCCTGTGAACCACAGCTGCACAGAAGG 2476
Qy 1989 CCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAA 2048
Db 2477 CCAGATCTCCATTGCCAAGTATGAGAGTTGCCCCAAGGATAACCCCATGTACTACTGCAA 2536
Qy 2049 CAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAA 2108
Db 2537 TAAGAAAACCAGCTGCAGGAGCTGTGCCCTAGACCAGAACTGCCAGTGGGAGCCCCGGAA 2596
Qy 2109 TCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAA 2168
Db 2597 TCAAGAGTGCATCGCCCTGCCGGAAAATATCTGTGGCAATGGCTGGCATTTGGTTGGAAA 2656
Qy 2169 CTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAG 2228
Db 2657 CTCGTGTCTGAAAATCACTACTGCTAAGGAGAATTATGACAATGCTAAATTGTCCTGTAG 2716
Qy 2229 GAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAA 2288
Db 2717 GAACCACAATGCCTTTTTGGCTTCCCTCACATCCCAGAAGAAGGTGGAGTTTGTCCTTAA 2776
Qy 2289 GCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGT 2348
Db 2777 GCAGCTTCGATTAATGCAATCATCTCAAAGTATGTCCAAGCTCACTCTGACTCCATGGGT 2836
Qy 2349 CGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAA 2408
Db 2837 TGGTCTTCGGAAGATCAATGTGTCTTACTGGTGCTGGGAGGATATGTCTCCATTCACAAA 2896
Qy 2409 TAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATC 2468
Db 2897 TAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGTGATGCTGGCTTCTGTGGGATCTTGTC 2956
Qy 2469 AGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGT 2528
Db 2957 AGAGCCTAGTACTCGGGGATTAAAGGCTGCAACCTGCATCAACCCTCTCAATGGCAGCGT 3016
Qy 2529 CTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGAC 2588
Db 3017 CTGTGAAAGGCCTGCAAACCACAGTGCCAAGCAGTGCCGGACACCATGTGCCCTGCGGAC 3076
Qy 2589 AGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCA 2648
Db 3077 AGCGTGTGGCGAGTGCACTAGCAGCAGCTCGGAGTGCATGTGGTGCAGTAACATGAAGCA 3136

```
Qy       2649 GTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTA 2708
             ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       3137 GTGTGTGGACTCCAATGCCTACGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTA 3196

Qy       2709 TACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTT 2768
             |||||||||||  |||||  |||||||||||  ||  |||||||||||||||  ||  ||||||||
Db       3197 TACGATGAGCAGCTGCCCACCTGAAAATTGCTCTGGCTACTGTACCTGCAGCCATTGCTT 3256

Qy       2769 GGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCAT 2828
             ||||||  ||||||||||||||||||||||||||||||  ||||||||||  |||||  ||
Db       3257 GGAGCAGCCAGGCTGTGGTTGGTGTACTGATCCTAGCAATACTGGGAAAGGAAAATGTAT 3316

Qy       2829 AGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTA 2888
               |||||   ||||||||||||  ||||||||||  ||  ||  |||  ||  ||||||||  | ||
Db       3317 TGAGGGCAGCTATAAAGGACCTGTGAAGATGCCGTCACAGGCCTCTGCAGGAAATGTGTA 3376

Qy       2889 TCCACAGCCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTT 2948
             ||||||||||||  ||  ||||||||||||||||||||||||||||||||||||||||||||
Db       3377 TCCACAGCCCCTTCTGAACTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTT 3436

Qy       2949 CATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTG 3008
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3437 CATTCACTGTCCAGCTTGCCAGTGCAACGGACACAGCAAATGCATCAACCAGAGTATCTG 3496

Qy       3009 TGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTA 3068
             ||||||||||||  ||||||||||  ||||||||||||||||||||||||||||||||||||
Db       3497 TGAGAAGTGTGAGGACCTGACCACGGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTA 3556

Qy       3069 CGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCT 3128
               |||||  ||  ||  ||||||||  ||||||||||||||||||||||||||||||  ||  ||
Db       3557 TGGTGACCCGACTAATGGAGGCAAATGTCAGCCATGCAAGTGCAATGGGCACGCATCACT 3616

Qy       3129 GTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTG 3188
             ||||||||||||||  ||||||||||||||  ||||||||  ||  ||||||||||||||||
Db       3617 GTGCAACACCAACACCGGCAAGTGCTTCTGTACCACCAAAGGTGTCAAGGGGGACGAGTG 3676

Qy       3189 CCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTA 3248
             ||||||||||||||||||||||||||||||||||||||||||||  ||||||||||  ||  ||
Db       3677 CCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAAAGGAACATGCTACTA 3736

Qy       3249 TACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTA 3308
               |||  |||||  ||||||||||||||||||  ||  ||||||||||  ||  |||||  ||
Db       3737 TACCCTTCTCATTGACTATCAGTTCACCTTTAGCCTGTCCCAGGAAGACGACCGCTACTA 3796

Qy       3309 CACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCAT 3368
             ||||||  |||||  ||||||||||||||||||  ||||||||||||||||||||||||||||
Db       3797 CACAGCCATCAACTTTGTGGCTACTCCTGATGAACAAAACAGGGATTTGGACATGTTCAT 3856

Qy       3369 CAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAAC 3428
             ||||||||||||  ||  ||||||||||||||||||||  ||||  |||  ||||  |||||
Db       3857 CAATGCCTCCAAAAACTTCAACCTCAACATCACCTGGGCCACCAGCTTCCCAGCCGGAAC 3916

Qy       3429 CCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAG 3488
             ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3917 CCAGACTGGAGAAGAGGTGCCTGTTGTTTCAAAAACCAACATCAAGGAATACAAAGATAG 3976

Qy       3489 TTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGT 3548
               |||||||||||||  ||||||||||||||||||  |||||  ||||||||||||||||||||
Db       3977 CTTCTCTAATGAGAAATTTGATTTTCGCAACCATCCAAACATCACTTTCTTTGTTTATGT 4036

Qy       3549 CAGTAATTTCACCTGGCCCATCAAAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTAT 3608
             ||||||||||||  ||||||||||||||||||||||||||||||  ||||||||||  ||  ||
Db       4037 CAGTAATTTCACTTGGCCCATCAAAATTCAGATTGCCTTCTCCCAGCACAGCAACTTCAT 4096

Qy       3609 GGACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGC 3668
             ||||||||||||||||||||||||||||||||||||||||||||||||||||  ||||||||
Db       4097 GGACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCGCTGCTTCTGGTGGC 4156

Qy       3669 TGCTGTGGTTTGGAAGATCAAACAAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCT 3728
             |||  |||||  ||||||||||  ||  ||  ||||||||  ||||  ||  ||||||||||||||
Db       4157 TGCAGTGGTCTGGAAGATCAAGCAGAGCTGTTGGGCATCCAGGCGGAGAGAGCAACTTCT 4216

Qy       3729 TCGAGAGATGCAACAGATGGCCAGCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAAC 3788
             |||  |||||||||||||||||||||||||||||||||  ||||||||  ||||||||  |||
Db       4217 TCGGGAGATGCAACAGATGGCCAGCCGCCCCTTTGCTTCTGTAAACGTTGCCTTGGAAAC 4276

Qy       3789 AGATGAGGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGC 3848
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4277 AGATGAGGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGC 4336

Qy       3849 ACTGGAGCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCG 3908
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4337 ACTGGAGCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCG 4396

Qy       3909 AGGCCTGGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGT 3968
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4397 AGGCCTGGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGT 4456

Qy       3969 GGACATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCG 4028
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4457 GGACATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCG 4516

Qy       4029 GAAGCAGCAGCCCCCTGCACAGCCTGGGACCTGCATCTGA 4068
             ||||||||||||||||||||||||||||||||||||||||
Db       4517 GAAGCAGCAGCCCCCTGCACAGCCTGGGACCTGCATCTGA 4556
```

RESULT 4
US-09-893-238-18
; Sequence 18, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 2625
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-893-238-18

Query Match 50.4%; Score 2051.8; DB 9; Length 2625;
Best Local Similarity 97.1%; Pred. No. 0;
Matches 2089; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

```
Qy         15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74
              || |||   |  ||  | ||  |  |  |  | | |  ||||  ||  ||  || ||
Db        234 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAA 293

Qy         75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134
              |    | |  | ||  ||||||||||||||||||||||||||||||||||||||||||||
Db        294 GGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 353

Qy        135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        354 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 413

Qy        195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 473

Qy        255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 533
```

```
Qy       315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374
Db       534 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 593

Qy       375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434
Db       594 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 653

Qy       435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494
Db       654 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 713

Qy       495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554
Db       714 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 773

Qy       555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614
Db       774 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 833

Qy       615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 674
Db       834 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 893

Qy       675 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 734
Db       894 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 953

Qy       735 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 794
Db       954 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 1013

Qy       795 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 854
Db      1014 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 1073

Qy       855 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 914
Db      1074 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 1133

Qy       915 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 974
Db      1134 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 1193

Qy       975 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1034
Db      1194 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1253

Qy      1035 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1094
Db      1254 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1313

Qy      1095 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1154
Db      1314 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1373

Qy      1155 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1214
Db      1374 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1433

Qy      1215 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1274
Db      1434 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1493

Qy      1275 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1334
Db      1494 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1553

Qy      1335 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1394
Db      1554 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1613

Qy      1395 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1454
Db      1614 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1673

Qy      1455 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1514
Db      1674 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1733

Qy      1515 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1574
Db      1734 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1793

Qy      1575 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGC 1634
Db      1794 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGC 1853

Qy      1635 AGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGA 1694
Db      1854 AGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGA 1913

Qy      1695 CATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGC 1754
Db      1914 CATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGC 1973

Qy      1755 AGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGC 1814
Db      1974 AGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGC 2033

Qy      1815 GCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCT 1874
Db      2034 GCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCT 2093

Qy      1875 TGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGA 1934
Db      2094 TGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGA 2153

Qy      1935 CTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGAT 1994
Db      2154 CTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGAT 2213

Qy      1995 CTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAA 2054
Db      2214 CTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAA 2273

Qy      2055 GACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGA 2114
Db      2274 GACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGA 2333

Qy      2115 GTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGG 2165
Db      2334 GTGCATTGCCCTGCCCGGTAGGCCTTGCAGGGTCATCTTGGTGTGTGTGGG 2384
```

RESULT 5
US-09-893-238-12
; Sequence 12, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 12
; LENGTH: 6370
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-893-238-12

Query Match 45.6%; Score 1853; DB 9; Length 6370;
Best Local Similarity 96.3%; Pred. No. 0;
Matches 1980; Conservative 7; Mismatches 47; Indels 23; Gaps 8;

```
Qy       2035 ATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAG 2094
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAG 60

Qy       2095 TGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGG 2154
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGG 120

Qy       2155 CATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCT 2214
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCT 180

Qy       2215 AAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTA 2274
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTA 240

Qy       2275 GAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACC 2334
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACC 300

Qy       2335 TTAACCCCATGGGTCGGCCTT---CGGAAGATCAATGTGTCCTACTGG-----TGCTGGG 2386
             |||||||||||||||||| | |   | | |:|||||||||||||||:|     |     |
Db        301 TTAACCCCATGGGTCGGGCCTTCGGGAAGGTYCAATGTGTCCTACTKGGTGCTGGGGAAG 360

Qy       2387 AAGATATGTCCCCATTTACAAATAG-TTTACTACAGTGGATGCC----GTCTGAGCCCAG 2441
              | || :  |||  |||||||||| |||||||||||    :     |    ||||||
Db        361 GATATGKTCCCCATTTTACAAATAGTTTTACTACAGTGGGATGSCCGTCTTGAGGCCCAG 420

Qy       2442 TGATGCT--GGATTCTGTGGAATTTT-----ATCAGAACCCAGTACTCGGGGACTGAAGG 2494
             ||  ||||  |:|||||||| | |||     | |   |         |||||||||||||
Db        421 TGTTGCTTGGRATTCTGTGGGAATTTTATTCAGGAACCCAGTTACTTCGGGGACTGAAGG 480

Qy       2495 CTGCAACCTGCAT--CAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAG 2552
             |||||||||||||  | ||   ::|||||||||||||||||||||||||||||||||||
Db        481 CTGCAACCTGCATTCAACCCACTYMAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAG 540

Qy       2553 TGCTAA-GCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCG 2611
             |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGCTAAGGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCG 600

Qy       2612 GCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATG 2671
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATG 660

Qy       2672 TGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTG 2731
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTG 720

Qy       2732 AAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGT 2791
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGT 780

Qy       2792 GTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAG 2851
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAG 840

Qy       2852 TGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCA 2911
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCA 900

Qy       2912 GCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAAT 2971
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAAT 960

Qy       2972 GCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCA 3031
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 GCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCA 1020

Qy       3032 CAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGA 3091
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 CAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGA 1080

Qy       3092 AATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGT 3151
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 AATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGT 1140

Qy       3152 GCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATC 3211
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 GCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATC 1200

Qy       3212 GATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGT 3271
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1201 GATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGT 1260

Qy       3272 TCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTA 3331
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1261 TCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTA 1320

Qy       3332 CTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACC 3391
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1321 CTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACC 1380

Qy       3392 TCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTG 3451
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1381 TCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTG 1440

Qy       3452 TTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATT 3511
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1441 TTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATT 1500

Qy       3512 TTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCA 3571
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1501 TTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCA 1560

Qy       3572 AAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCTGGTACAGTTCTTCGTGA 3631
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1561 AAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCTGGTACAGTTCTTCGTGA 1620

Qy       3632 CTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGTGGTTTGGAAGATCAAAC 3691
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1621 CTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGTGGTTTGGAAGATCAAAC 1680

Qy       3692 AAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGAGATGCAACAGATGGCCA 3751
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1681 AAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGAGATGCAACAGATGGCCA 1740

Qy       3752 GCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTA 3811
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1741 GCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTA 1800

Qy       3812 TTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGAGCCGTGTTTTGGCAACA 3871
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1801 TTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGAGCCGTGTTTTGGCAACA 1860

Qy       3872 AAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCTGGGTGGCATCCCTCCTC 3931
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1861 AAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCTGGGTGGCATCCCTCCTC 1920

Qy       3932 CTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGA 3991
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1921 CTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGA 1980

Qy       3992 TAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGC 4051
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1981 TAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGC 2040
```

```
Qy       4052 CTGGGACCTGCATCTGA 4068
              |||||||||||||||||
Db       2041 CTGGGACCTGCATCTGA 2057


RESULT 6
US-10-197-824-1
; Sequence 1, Application US/10197824
; Publication No. US20040023219A1
; GENERAL INFORMATION:
;  APPLICANT: ORIGENE TECHNOLOGIES INC
;  TITLE OF INVENTION: NOVEL PROSTATE CANCER GENES
;  FILE REFERENCE: 9U 106 R1
;  CURRENT APPLICATION NUMBER: US/10/197,824
;  CURRENT FILING DATE:  2002-07-19
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 6733
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (575)..(4711)
;   OTHER INFORMATION:
US-10-197-824-1

  Query Match           35.9%;  Score 1461.8;  DB 16;  Length 6733;
  Best Local Similarity   63.1%;  Pred. No. 0;
  Matches 2495;  Conservative    0;  Mismatches 1372;  Indels   90;  Gaps   12;

Qy        110 GCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGC 169
              ||  |  || || |  |   ||| |||||  | | ||||||||||| || || |||||||
Db        789 GCCGCTGTGTCAACTCCACCTGCCTCTGCGACCCGGGCTGGGTGGGGGACCAGTGCCAGC 848

Qy        170 ACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAA 229
              |||||  |||| | ||||  |||| | | |||||||| | | | ||||||||  ||    |
Db        849 ACTGCCAGGGCAGGTTCAAGTTAACAGAACCTTCTGGATATTTAACAGATGGCCCAATTA 908

Qy        230 ATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGA 289
              | |||||||| ||||| || || || ||||||||||||||  | ||||||  | |  | |
Db        909 ACTATAAATATAAAACTAAATGTACTTGGCTCATTGAAGGCTATCCAAATGCAGTGTTAA 968

Qy        290 GACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATG 349
              || |  | |||||||||||||||||||| ||||| ||||| ||| | ||||||||||||
Db        969 GATTAAGATTCAATCATTTTGCTACAGAATGTAGCTGGGATCATATGTATGTTTATGATG 1028

Qy        350 GGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAG 409
              | || ||||| |||||||| || | |||| | |||||||  | || || ||||| | | |
Db       1029 GAGATTCAATATATGCACCTTTAATAGCTGTACTTAGTGGTTTGATAGTCCCTGAAATAA 1088

Qy        410 ATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTT 469
              ||||||||  |||||  |||||  |||||| | |||||  || |||||   || | |||||||
Db       1089 GGGGCAATGAAACTGTGCCTGAAGTTGTTACTACATCTGGCTATGCACTGTTACATTTTT 1148

Qy        470 TTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTC 529
              |||||||||||||  ||||||  | |||||  || ||  |||   ||     ||   ||||
Db       1149 TTAGTGATGCTGCGTATAATCTAACTGGTTTCAACATTTTCTATTCAATCAATTCTTGTC 1208

Qy        530 CAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGT---AATAGCAGCGATACTGTTG 586
              | || || ||||| || |  ||  ||||||  |  |||  |     ||  |  |  |  ||
Db       1209 CTAACAATTGCTCTGGTCATGGGAAGTGTACAACTAGTGTCTCTGTTCCAAGTCAAGTAT 1268

Qy        587 AATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACA 646
              |  |||||||||   | ||  |||||||  |||||||  |||||  ||||||  |||| || ||
Db       1269 ATTGTGAATGTGATAAATACTGGAAGGGTGAAGCTTGTGATATTCCTTACTGTAAAGCCA 1328

Qy        647 ACTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCT 706
              | || ||  |||  |||  ||  |||  |  | ||  |  |  |  |||  ||||  ||||
Db       1329 ATTGCGGCAGTCCAGATCACGGTTACTGTGACCTGACTGGAGAAAAATTATGTGTCTGCA 1388

Qy        707 TCTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGA 766
                  |||||  |||||||  |||||  |  |||  ||  |||  ||||  |  ||||

Db       1389 ATGATAGTTGGCAAGGTCCTGATTGTTCTTTGAATGTTCCCTCTACTGAGTCTTACTGGA 1448

Qy        767 CTCGAGAGGAATATTCTAACTTAAAGC------TCCCCAGAGCATCTCATAAAGCTGTGG 820
              ||       |   |  |||  |   |     |  || || ||||||||  ||
Db       1449 TTCTGCCAAACGTTAAACCCTTCAGTCCTTCTGTAGGTCGGGCTTCACATAAAGCAGTTT 1508

Qy        821 TCAATGGAAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACA 880
              |  | || ||  ||||||||||  |||| ||||||| || ||| ||     || | | |
Db       1509 TACACGGGAAATTTATGTGGGTGATTGGTGGATATACTTTTAACTACAGTTCTTTTCAAA 1568

Qy        881 TGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACA 940
              |||| |||   ||  |  | |    ||   |||  |  |  ||  |  ||  |
Db       1569 TGGTCCTAAATTACAATTTAGAAAGCAGTATATGGAATGTAGGAACTCCATCAAGGGGAC 1628

Qy        941 ATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACATGTATG 1000
              |  |  ||||||||  ||  |||  |  || |||||  ||||  ||  ||  |  |||||||
Db       1629 CT---CTCCAGAGATATGGACACTCTCTTGCTTTATATCAGGAAAACATCTTTATGTATG 1685

Qy       1001 GAGGAAAAATTGATTCAA---CTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTC 1057
              ||||  |  ||||||  |||    ||| |||||  ||  ||||  ||  |  |||||| |||| |
Db       1686 GAGGCAGAATTGAAACAAATGATGGCAATGTCACAGATGAATTATGGGTTTTTAACATAC 1745

Qy       1058 ATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGG---------AGCAGTATGCAG 1108
              |||  | ||||||||           ||  ||||  |       |           |||||||||| |
Db       1746 ATAGTCAGTCATGGAGTACAAAAACTCCTACTGTTCTTGGACATGGTCAGCAGTATGCTG 1805

Qy       1109 TGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCA 1168
              |||| || || ||||  |||  |       || | | | |    || |||||  | | ||
Db       1806 TGGAGGGACATTCAGCACATATTATGGAGTTGGATAGTAGAGATGTTGTCATGATCATAA 1865

Qy       1169 TCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATA 1228
              | |||||  |  |   | |||||  |||| ||||  |  ||||||||  || |
Db       1866 TATTTGGATATTCTGCAATATATGGTTATACAAGCAGCATACAGGAATACCATATCTCAT 1925

Qy       1229 AGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATA 1288
              |||||  |||  | |  | | ||  | || ||  |||| |||||  || || |||||||
Db       1926 CAAACACTTGGCTTGTTCCAGAAACTAAAGGAGCTATTGTACAAGGTGGATATGGCCATA 1985

Qy       1289 GCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCA 1348
              |||||  || ||  | | || | || | || | ||||||||  || || || || ||  |
Db       1986 CTAGTGTGTATGATGAAATAACAAAGTCCATTTATGTTCATGGAGGGTATAAAGCATTGC 2045

Qy       1349 GTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGT 1408
              |  || || ||  |  | |  |||||||| ||  |||||  ||  | ||  |||  |
Db       2046 CAGGGAACAAATATGGATTGGTTGATGATCTTTATAAATATGAAGTTAACACTAAGACTT 2105

Qy       1409 GGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTG 1468
              |||| ||| |  || ||  ||  | |||  | |  ||  |||||||  | | ||
Db       2106 GGACTATTTTGAAAGAAAGTGGGTTTGCCAGATACCTTCATTCAGCTGTTCTTATCAATG 2165

Qy       1469 GAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGCCATGGCGCCA 1528
              || | |||||  | |||||  |||||  || || ||||||||  ||  ||||  | || || |
Db       2166 GAGCTATGCTTATTTTTGGAGGAAATACCCATAATGACACTTCCTTGAGTAACGGTGCAA 2225

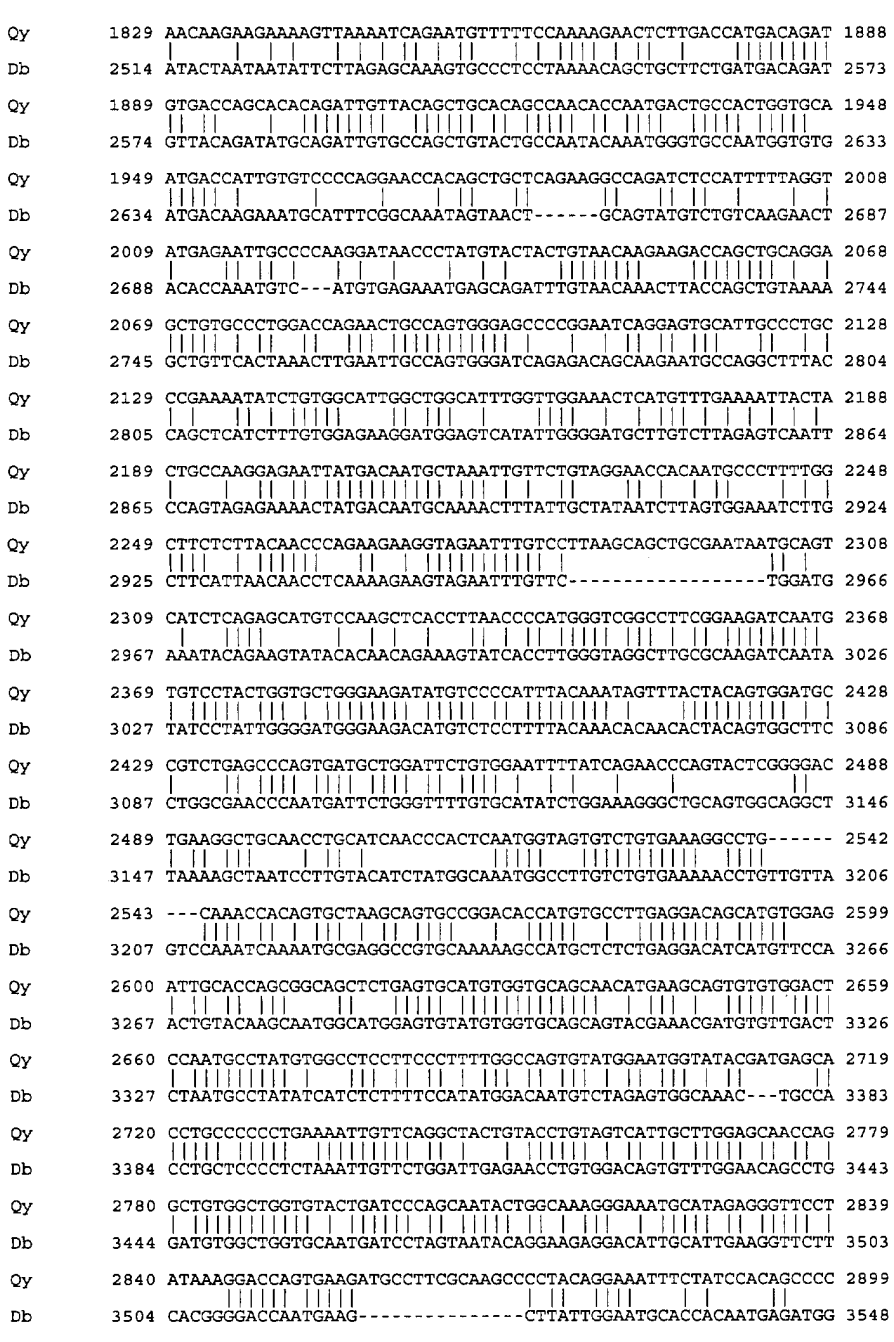
Qy       1529 AATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTC 1588
              |||| || ||| | |||||| |||| ||||||||  || |||||     |||  | | || |
Db       2226 AATGTTTTTCTGCCGATTTCCTGGCATATGACATAGCTTGTGATGAATGGAAAATACTAC 2285

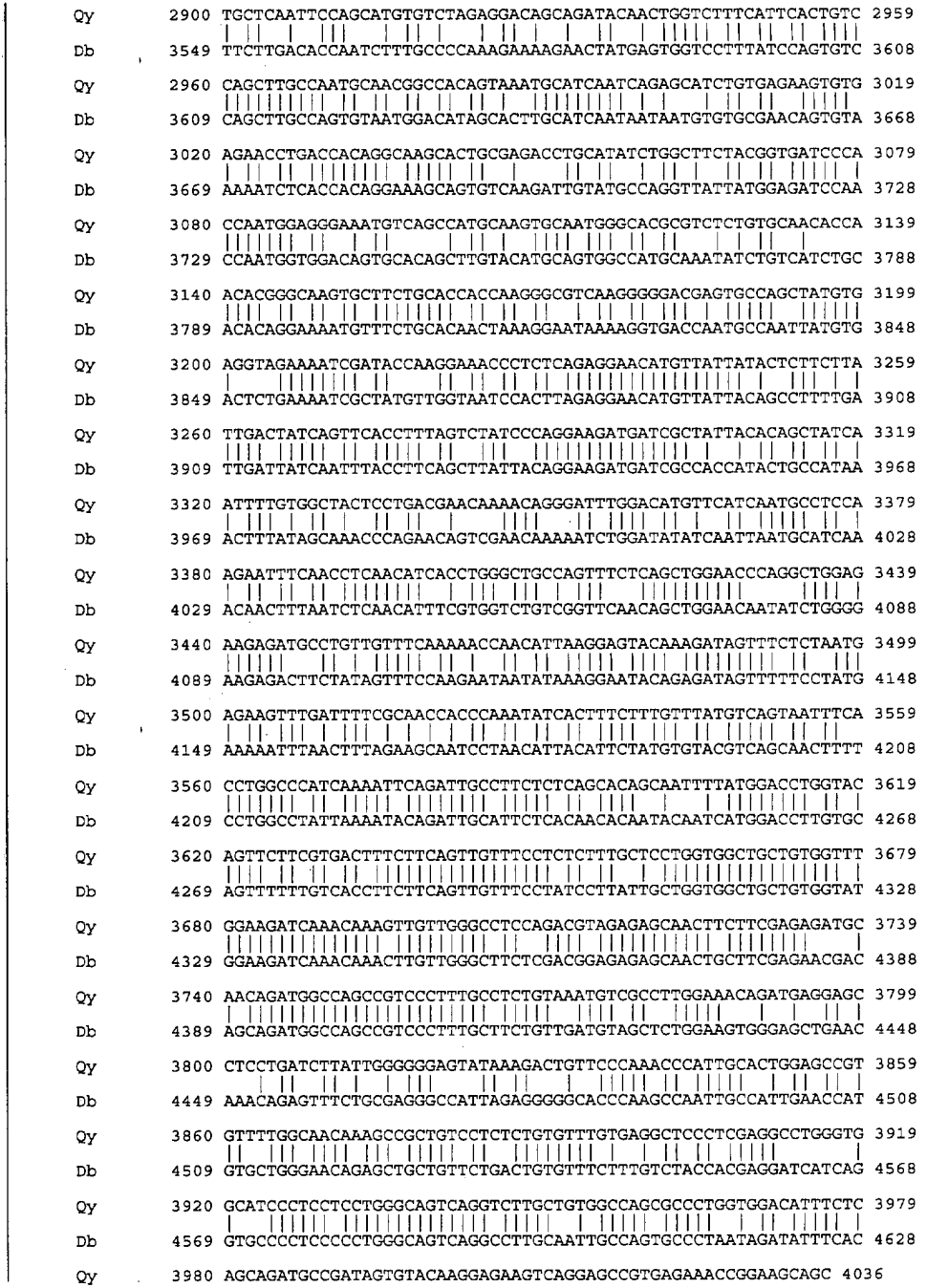
Qy       1589 CCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACAACA 1648
              | | |||  |||| ||  |||||||||||||||||  || || |||||  |  |||
Db       2286 CAAAACCAAATCTTCATAGAGATGTCAACAGATTTGGACACTCTGCAGTAGTCATTAACG 2345

Qy       1649 GCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCA 1708
              |  |||||||  | || || || ||  |||| | |||||  | || |||||  |||| ||
Db       2346 GGTCCATGTATATATTTGGGGGATTTTCTAGTGTACTCCTTAATGATATCCTTGTATACA 2405

Qy       1709 CCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTA 1768
              |  | | | || |  | ||  |  ||||  |  || || || || || ||  ||| ||
Db       2406 AGCCTCCAAATTGCAAGGCTTTCAGAGATGAAGAACTTTGTAAAAATGCTGGTCCAGGGA 2465

Qy       1769 TTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATG 1828
              |  |||| ||||| | |  | | | |                  | || | |||
Db       2466 TAAAATGTGTTTGGAATAAAAATCACTGTGAAT-----------CTTGGGAATCTGGGA 2513
```


Qy 1829 AACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGAT 1888
Db 2514 ATACTAATAATATTCTTAGAGCAAAGTGCCCTCCTAAAACAGCTGCTTCTGATGACAGAT 2573
Qy 1889 GTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGCA 1948
Db 2574 GTTACAGATATGCAGATTGTGCCAGCTGTACTGCCAATACAAATGGGTGCCAATGGTGTG 2633
Qy 1949 ATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGT 2008
Db 2634 ATGACAAGAAATGCATTTCGGCAAATAGTAACT------GCAGTATGTCTGTCAAGAACT 2687
Qy 2009 ATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGGA 2068
Db 2688 ACACCAAATGTC---ATGTGAGAAATGAGCAGATTTGTAACAAACTTACCAGCTGTAAAA 2744
Qy 2069 GCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGC 2128
Db 2745 GCTGTTCACTAAACTTGAATTGCCAGTGGGATCAGAGACAGCAAGAATGCCAGGCTTTAC 2804
Qy 2129 CCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTA 2188
Db 2805 CAGCTCATCTTTGTGGAGAAGGATGGAGTCATATTGGGGATGCTTGTCTTAGAGTCAATT 2864
Qy 2189 CTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGG 2248
Db 2865 CCAGTAGAGAAAACTATGACAATGCAAAACTTTATTGCTATAATCTTAGTGGAAATCTTG 2924
Qy 2249 CTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGT 2308
Db 2925 CTTCATTAACAACCTCAAAAGAAGTAGAATTTGTTC------------------TGGATG 2966
Qy 2309 CATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATG 2368
Db 2967 AAATACAGAAGTATACACAACAGAAAGTATCACCTTGGGTAGGCTTGCGCAAGATCAATA 3026
Qy 2369 TGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGC 2428
Db 3027 TATCCTATTGGGGATGGGAAGACATGTCTCCTTTTACAAACACAACACTACAGTGGCTTC 3086
Qy 2429 CGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGAC 2488
Db 3087 CTGGCGAACCCAATGATTCTGGGTTTTGTGCATATCTGGAAAGGGCTGCAGTGGCAGGCT 3146
Qy 2489 TGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTG------ 2542
Db 3147 TAAAAGCTAATCCTTGTACATCTATGGCAAATGGCCTTGTCTGTGAAAAACCTGTTGTTA 3206
Qy 2543 ---CAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAG 2599
Db 3207 GTCCAAATCAAAATGCGAGGCCGTGCAAAAAGCCATGCTCTCTGAGGACATCATGTTCCA 3266
Qy 2600 ATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACT 2659
Db 3267 ACTGTACAAGCAATGGCATGGAGTGTATGTGGTGCAGCAGTACGAAACGATGTGTTGACT 3326
Qy 2660 CCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCA 2719
Db 3327 CTAATGCCTATATCATCTCTTTTCCATATGGACAATGTCTAGAGTGGCAAAC---TGCCA 3383
Qy 2720 CCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAG 2779
Db 3384 CCTGCTCCCCTCTAAATTGTTCTGGATTGAGAACCTGTGGACAGTGTTTGGAACAGCCTG 3443
Qy 2780 GCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCT 2839
Db 3444 GATGTGGCTGGTGCAATGATCCTAGTAATACAGGAAGAGGACATTGCATTGAAGGTTCTT 3503
Qy 2840 ATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCC 2899
Db 3504 CACGGGGACCAATGAAG--------------CTTATTGGAATGCACCACAATGAGATGG 3548


Qy 2900 TGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTC 2959
Db 3549 TTCTTGACACCAATCTTTGCCCCAAAGAAAAGAACTATGAGTGGTCCTTTATCCAGTGTC 3608
Qy 2960 CAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTG 3019
Db 3609 CAGCTTGCCAGTGTAATGGACATAGCACTTGCATCAATAATAATGTGTGCGAACAGTGTA 3668
Qy 3020 AGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCA 3079
Db 3669 AAAATCTCACCACAGGAAAGCAGTGTCAAGATTGTATGCCAGGTTATTATGGAGATCCAA 3728
Qy 3080 CCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCA 3139
Db 3729 CCAATGGTGGACAGTGCACAGCTTGTACATGCAGTGGCCATGCAAATATCTGTCATCTGC 3788
Qy 3140 ACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTG 3199
Db 3789 ACACAGGAAAATGTTTCTGCACAACTAAAGGAATAAAAGGTGACCAATGCCAATTATGTG 3848
Qy 3200 AGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTA 3259
Db 3849 ACTCTGAAAATCGCTATGTTGGTAATCCACTTAGAGGAACATGTTATTACAGCCTTTTGA 3908
Qy 3260 TTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCA 3319
Db 3909 TTGATTATCAATTTACCTTCAGCTTATTACAGGAAGATGATCGCCACCATACTGCCATAA 3968
Qy 3320 ATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCA 3379
Db 3969 ACTTTATAGCAAACCCAGAACAGTCGAACAAAAATCTGGATATATCAATTAATGCATCAA 4028
Qy 3380 AGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAG 3439
Db 4029 ACAACTTTAATCTCAACATTTCGTGGTCTGTCGGTTCAACAGCTGGAACAATATCTGGGG 4088
Qy 3440 AAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATG 3499
Db 4089 AAGAGACTTCTATAGTTTCCAAGAATAATATAAAGGAATACAGAGATAGTTTTTCCTATG 4148
Qy 3500 AGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCA 3559
Db 4149 AAAAATTTAACTTTAGAAGCAATCCTAACATTACATTCTATGTGTACGTCAGCAACTTTT 4208
Qy 3560 CCTGGCCCATCAAAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCTGGTAC 3619
Db 4209 CCTGGCCTATTAAAATACAGATTGCATTCTCACAACACAATACAATCATGGACCTTGTGC 4268
Qy 3620 AGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGTGGTTT 3679
Db 4269 AGTTTTTTGTCACCTTCTTCAGTTGTTTCCTATCCTTATTGCTGGTGGCTGCTGTGGTAT 4328
Qy 3680 GGAAGATCAAACAAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGAGATGC 3739
Db 4329 GGAAGATCAAACAAACTTGTTGGGCTTCTCGACGGAGAGAGCAACTGCTTCGAGAACGAC 4388
Qy 3740 AACAGATGGCCAGCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGAGGAGC 3799
Db 4389 AGCAGATGGCCAGCCGTCCCTTTGCTTCTGTTGATGTAGCTCTGGAAGTGGGAGCTGAAC 4448
Qy 3800 CTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGAGCCGT 3859
Db 4449 AAACAGAGTTTCTGCGAGGGCCATTAGAGGGGGCACCCAAGCCAATTGCCATTGAACCAT 4508
Qy 3860 GTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCTGGGTG 3919
Db 4509 GTGCTGGGAACAGAGCTGCTGTTCTGACTGTGTTTCTTTGTCTACCACGAGGATCATCAG 4568
Qy 3920 GCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACATTTCTC 3979
Db 4569 GTGCCCCTCCCCCTGGGCAGTCAGGCCTTGCAATTGCCAGTGCCCTAATAGATATTTCAC 4628
Qy 3980 AGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCAGC 4036

```
Db       4629 AACAGAAAGCTTCAGATAGTAAAGATAAGACTTCTGGAGTCCGGAATCGAAAACACC 4685


RESULT 7
US-09-893-238-8
; Sequence 8, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT:  Nagle, D.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
;  FILE REFERENCE: 7853-237
;  CURRENT APPLICATION NUMBER: US/09/893,238
;  CURRENT FILING DATE:  2001-06-27
;  PRIOR APPLICATION NUMBER: 09/245,041
;  PRIOR FILING DATE: 1999-02-05
;  PRIOR APPLICATION NUMBER: 60/093,630
;  PRIOR FILING DATE: 1998-07-21
;  PRIOR APPLICATION NUMBER: 60/104,978
;  PRIOR FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
;   LENGTH: 2419
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-893-238-8

  Query Match             35.7%;  Score 1451.8;  DB 9;  Length 2419;
  Best Local Similarity   89.6%;  Pred. No. 0;
  Matches 1561;  Conservative    0;  Mismatches  182;  Indels    0;  Gaps    0;

Qy        189 ACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAA 248
Db        677 ACTAACTGGCTCTTCTGGATTTGTAACAGATGGACCTGGGAATTATAAATATAAGACGAA 736

Qy        249 GTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTT 308
Db        737 GTGCACATGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGCTTCAACCATTT 796

Qy        309 TGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACC 368
Db        797 TGCTACAGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATCTACGCACC 856

Qy        369 GCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCC 428
Db        857 TCTGATTGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCTCC 916

Qy        429 TGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAA 488
Db        917 TGAGGTCACTGTCACTTCAGGTTATGCACTGCTGCATTTTTTCAGTGATGCTGCTTATAA 976

Qy        489 TTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCG 548
Db        977 TCTGACTGGATTTAATATCACTTACAATTTTGACATGTGTCCGAATAATTGCTCAGGCCG 1036

Qy        549 AGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTG 608
Db       1037 AGGAGAGTGTAAGAGCAGTAACAGCAGCAGCGCTGTTGAGTGTGAATGTTCTGAAAACTG 1096

Qy        609 GAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGG 668
Db       1097 GAAAGGGGAGTCGTGTGACATTCCTCACTGTACAGACAACTGTGGCTTTCCTCACCGAGG 1156

Qy        669 CATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGG 728
Db       1157 CATCTGTAATGCAAGCGATACCAGAGGGTGCTCCTGCTTTCCTCACTGGCAGGGTCCTGG 1216

Qy        729 ATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTT 788
Db       1217 ATGTTCAATTCCTGTGCCAGCTAACCAGTCTTTTTGGACTCGAGAAGAATATTCTGATTT 1276

Qy        789 AAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGG 848
Db       1277 AAAGCTTCCCAGAGCCTCTCATAAAGCTGTGGTCAATGGAAATATAATGTGGGTTGTTGG 1336

Qy        849 AGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAG 908
Db       1337 CGGATATATGTTCAACCATTCAGATTACAGCATGGTTTTAGCGTATGACCTGACTTCTAG 1396

Qy        909 GGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTT 968
Db       1397 GGAATGGCTTCCACTAAACCATTCTGTGAACAGTGTGGTTGTAAGATATGGTCATTCTTT 1456

Qy        969 GGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGT 1028
Db       1457 GGCATTACATAAGGATAAAATCTACATGTATGGAGGAAAAATTGATTCAACAGGGAACGT 1516

Qy       1029 GACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAA 1088
Db       1517 GACCAATGAGCTGAGAGTATTTCATATTCATAATGAATCATGGGTATTGTTAACTCCGAA 1576

Qy       1089 GGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGG 1148
Db       1577 AGCTAAGGATCAGTATGCAGTGGTTGGACACTCAGCACACATTGTTACACTGGCATCTGG 1636

Qy       1149 CCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGT 1208
Db       1637 CCGTGTGGTCATGTTGGTCATCTTCGGTCATTGCCCACTCTATGGATATATAAGCGTTGT 1696

Qy       1209 GCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGT 1268
Db       1697 GCAGGAATATGACTTGGAAAAGAACACATGGAGTATATTACATACTCAGGGTGCTCTTGT 1756

Qy       1269 GCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCA 1328
Db       1757 GCAAGGGGGTTATGGCCACAGTAGTGTTTATGATGACAGGACCAAGGCTCTGTACGTTCA 1816

Qy       1329 TGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATA 1388
Db       1817 TGGTGGCTACAAGGCTTTCAGCGCCAACAAATACCGGCTTGCAGATGACCTCTACAGATA 1876

Qy       1389 TGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCA 1448
Db       1877 CGATGTGGATACTCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCA 1936

Qy       1449 CACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACAC 1508
Db       1937 TACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGAGGGAACACACACAATGACAC 1996

Qy       1509 ATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTG 1568
Db       1997 TTCCATGAGCCACGGTGCCAAATGCTTCTCCTCGGACTTCATGGCTTATGACATTGCTTG 2056

Qy       1569 TGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCA 1628
Db       2057 TGACCGATGGTCAGTGCTTCCCAGACCTGAGCTCCATCATGATGTCAACAGATTTGGCCA 2116

Qy       1629 TTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCT 1688
Db       2117 TTCAGCAGTCTTGTACAACAGCACCATGTATGTGTTCGGCGGCTTCAACAGCCTCCTCCT 2176

Qy       1689 CAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTG 1748
Db       2177 CAGTGACGTCTTGGTCTTTACCTCGGAGCAGTGCGATGCACACCGCAGTGAAGCTGCTTG 2236

Qy       1749 TTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTC 1808
Db       2237 TGTGGCAGCAGGACCTGGTATCCGGTGTCTGTGGGACACACAGTCGTCTCGATGTACCTC 2296

Qy       1809 GTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAG 1868
Db       2297 CTGGGAGTTGGCAACTGAAGAACAAGCAGAAAAGTTAAAATCAGAGTGTTTTTCTAAAAG 2356
```

```
Qy       1869 AACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACAC 1928
Db       2357 AACCCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAATAC 2416

Qy       1929 CAA 1931
Db       2417 CAA 2419
```

RESULT 8
US-09-893-238-10
; Sequence 10, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 1051
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-893-238-10

Query Match 11.4%; Score 463; DB 9; Length 1051;
Best Local Similarity 85.2%; Pred. No. 2.4e-132;
Matches 517; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

```
Qy         15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74
Db        441 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAA 500

Qy         75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134
Db        501 GGAATGTGACCGGCCGTGTGTCAACGGCGGCCGCTGCAACCCTGGCACCGGCCAGTGCGT 560

Qy        135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194
Db        561 CTGCCCCACGGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 620

Qy        195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254
Db        621 TGGCTCTTCTGGATTTGTAACAGATGGACCTGGGAATTATAAATATAAGACGAAGTGCAC 680

Qy        255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314
Db        681 ATGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGCTTCAACCATTTTGCTAC 740

Qy        315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374
Db        741 AGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATCTACGCACCTCTGAT 800

Qy        375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434
Db        801 TGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCTCCTGAGGT 860

Qy        435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494
Db        861 CACTGTCACTTCAGGTTATGCACTGCTGCATTTTTTCAGTGATGCTGCTTATAATCTGAC 920

Qy        495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554
```

```
Db        921 TGGATTTAATATCACTTACAATTTTGACATGTGTCCGAATAATTGCTCAGGCCGAGGAGA 980

Qy        555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614
Db        981 GTGTAAGAGCAGTAACAGCAGCAGCGCTGTTGAGTGTGAATGTTCTGAAAACTGGAAAGG 1040

Qy        615 TGAAGCA 621
Db       1041 GGCCGGA 1047
```

RESULT 9
US-10-198-846-5869/c
; Sequence 5869, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5869
; LENGTH: 625
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3, 4, 5, 6, 7, 8, 10, 12, 14, 15, 17, 26, 28, 29
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-5869

Query Match 6.8%; Score 277.4; DB 14; Length 625;
Best Local Similarity 99.6%; Pred. No. 9.5e-75;
Matches 278; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       2477 GTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAA 2536
Db        340 GTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAA 281

Qy       2537 GGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTG 2596
Db        280 GGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTG 221

Qy       2597 GAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGG 2656
Db        220 GAGATTGCACCAGCGGCAGCTCTGAGTGTATGTGGTGCAGCAACATGAAGCAGTGTGTGG 161

Qy       2657 ACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGA 2716
Db        160 ACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGA 101

Qy       2717 GCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCT 2755
Db        100 GCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCT 62
```

RESULT 10
US-09-864-761-15900
; Sequence 15900, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.

```
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 15900
;   LENGTH: 531
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL132773.14
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
US-09-864-761-15900

  Query Match             6.2%;  Score 250.8;  DB 9;  Length 531;
  Best Local Similarity   97.3%;  Pred. No. 1.6e-66;
  Matches  255;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy        973 TTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACC 1032
              || | | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 TTTTCCTAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACC 200

Qy       1033 AATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCA 1092
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 AATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCA 260

Qy       1093 AAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGA 1152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 AAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGA 320

Qy       1153 GTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAG 1212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 GTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAG 380

Qy       1213 GAATATGATTTGGATAAGAACA 1234
              ||||||||||||| || | | |
Db        381 GAATATGATTTGGGTAGGTATA 402


RESULT 11
US-09-736-457-966/c
; Sequence 966, Application US/09736457
; Patent No. US20020168637A1
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Wang, Aijun
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.478C15
;  CURRENT APPLICATION NUMBER: US/09/736,457
;  CURRENT FILING DATE:  2000-12-13
;  NUMBER OF SEQ ID NOS:  1864
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 966
;   LENGTH: 246
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-736-457-966

  Query Match             6.0%;  Score 246;  DB 9;  Length 246;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-65;
  Matches  246;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2293 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 2352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy       2353 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 2412
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy       2413 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 2472
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67

Qy       2473 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 2532
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy       2533 GAAAGG 2538
              ||||||
Db           6 GAAAGG 1


RESULT 12
US-09-902-941-966/c
; Sequence 966, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
;  APPLICANT: Henderson, Robert A.
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Watanabe, Yoshihiro
;  APPLICANT:  Johnson, Jeffrey C.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Marnerakis, Margarita
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Vedvick, Thomas S.
```

```
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  McNabb, Andria
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.478C17
;  CURRENT APPLICATION NUMBER: US/09/902,941
;  CURRENT FILING DATE:  2001-07-10
;  NUMBER OF SEQ ID NOS: 2002
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 966
;   LENGTH: 246
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-902-941-966

  Query Match             6.0%;  Score 246;  DB 9;  Length 246;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-65;
  Matches  246;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2293 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 2352
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy       2353 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 2412
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy       2413 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 2472
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67

Qy       2473 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 2532
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy       2533 GAAAGG 2538
             ||||||
Db            6 GAAAGG 1


RESULT 13
US-09-849-626-966/c
; Sequence 966, Application US/09849626
; Publication No. US20020197669A1
; GENERAL INFORMATION:
;  APPLICANT: Bangur, Chaitanya
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Switzer, Anne
;  APPLICANT:  McNeill, Patricia
;  APPLICANT:  Clapper, Jonathan
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.478C16
;  CURRENT APPLICATION NUMBER: US/09/849,626
;  CURRENT FILING DATE:  2001-05-03
;  NUMBER OF SEQ ID NOS: 1926
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 966
;   LENGTH: 246
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-849-626-966

  Query Match             6.0%;  Score 246;  DB 9;  Length 246;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-65;
  Matches  246;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2293 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 2352
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy       2353 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 2412
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy       2413 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 2472
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67

Qy       2473 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 2532
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy       2533 GAAAGG 2538
             ||||||
Db            6 GAAAGG 1


RESULT 14
US-10-017-754-966/c
; Sequence 966, Application US/10017754
; Publication No. US20030054363A1
; GENERAL INFORMATION:
;  APPLICANT: Henderson, Robert A.
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Watanabe, Yoshihiro
;  APPLICANT:  Johnson, Jeffrey C.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Marnerakis, Margarita
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  McNabb, Andria
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.478C18
;  CURRENT APPLICATION NUMBER: US/10/017,754
;  CURRENT FILING DATE:  2001-10-29
;  NUMBER OF SEQ ID NOS: 2004
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 966
;   LENGTH: 246
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-017-754-966

  Query Match             6.0%;  Score 246;  DB 14;  Length 246;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-65;
  Matches  246;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2293 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 2352
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy       2353 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 2412
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy       2413 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 2472
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67

Qy       2473 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 2532
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy       2533 GAAAGG 2538
             ||||||
Db            6 GAAAGG 1


RESULT 15
US-10-113-872-966/c
```

; Sequence 966, Application US/10113872
; Publication No. US20030170255A1
; GENERAL INFORMATION:
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C19
; CURRENT APPLICATION NUMBER: US/10/113,872
; CURRENT FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-113-872-966

Query Match 6.0%; Score 246; DB 14; Length 246;
Best Local Similarity 100.0%; Pred. No. 2.7e-65;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2293 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 2352
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy       2353 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 2412
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy       2413 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 2472
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67

Qy       2473 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 2532
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy       2533 GAAAGG 2538
             ||||||
Db           6 GAAAGG 1
```

Search completed: March 5, 2004, 13:18:32
Job time : 906.264 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2004, 17:02:19 ; Search time 180.783 Seconds
(without alignments)
12487.583 Million cell updates/sec

Title: US-09-787-097-11
Perfect score: 4068
Sequence: 1 atggtggccgcagcggcggc.....agcctgggacctgcatctga 4068

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	3988.4	98.0	8589	3	US-09-245-041-14	Sequence 14, Appl
	2	3499.8	86.0	4072	3	US-09-245-041-16	Sequence 16, Appl
	3	3326.4	81.8	8827	3	US-09-245-041-1	Sequence 1, Appli
	4	2051.8	50.4	2625	3	US-09-245-041-18	Sequence 18, Appl
	5	1853	45.6	6370	3	US-09-245-041-12	Sequence 12, Appl
	6	1451.8	35.7	2419	3	US-09-245-041-8	Sequence 8, Appli
	7	463	11.4	1051	3	US-09-245-041-10	Sequence 10, Appl
c	8	246	6.0	246	4	US-09-702-705-966	Sequence 966, App
c	9	246	6.0	246	4	US-09-736-457-966	Sequence 966, App
c	10	246	6.0	246	4	US-09-614-124B-966	Sequence 966, App
c	11	246	6.0	246	4	US-09-671-325-966	Sequence 966, App
	12	212.8	5.2	415	4	US-09-833-381-1916	Sequence 1916, Ap
	13	202.8	5.0	90050	3	US-09-245-041-5	Sequence 5, Appli
c	14	169.2	4.2	17056	3	US-09-245-041-3	Sequence 3, Appli
c	15	132.2	3.2	226	3	US-09-245-041-6	Sequence 6, Appli
c	16	115.4	2.8	221	3	US-09-245-041-7	Sequence 7, Appli
	17	100.2	2.5	5973	3	US-09-245-041-4	Sequence 4, Appli
c	18	48.8	1.2	4403765	3	US-09-103-840A-2	Sequence 2, Appli
c	19	48.8	1.2	4411529	3	US-09-103-840A-1	Sequence 1, Appli
	20	48.6	1.2	832	4	US-09-621-976-2813	Sequence 2813, Ap
	21	46.8	1.2	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c	22	46	1.1	2561	4	US-09-616-289-48	Sequence 48, Appl
	23	45.8	1.1	4403765	3	US-09-103-840A-2	Sequence 2, Appli
	24	44.8	1.1	5329	4	US-09-562-702A-19	Sequence 19, Appl
	25	44.8	1.1	5329	4	US-09-561-818A-19	Sequence 19, Appl
	26	44.8	1.1	5689	4	US-09-562-702A-17	Sequence 17, Appl
	27	44.8	1.1	5689	4	US-09-561-818A-17	Sequence 17, Appl
	28	42.6	1.0	4362	2	US-08-455-073A-1	Sequence 1, Appli
	29	42.4	1.0	700	4	US-09-236-097-7	Sequence 7, Appli
c	30	42	1.0	1213	1	US-08-276-151-6	Sequence 6, Appli
	31	41.6	1.0	538	4	US-09-056-556-180	Sequence 180, App
	32	41.6	1.0	538	4	US-09-072-596-175	Sequence 175, App
	33	41.6	1.0	538	4	US-09-072-967-180	Sequence 180, App
	34	41.4	1.0	1028	3	US-08-118-200-1	Sequence 1, Appli
	35	41.4	1.0	1028	3	US-08-458-745-1	Sequence 1, Appli
	36	41.4	1.0	3789	4	US-09-091-889A-3	Sequence 3, Appli
	37	41.4	1.0	8450	4	US-09-091-889A-5	Sequence 5, Appli
	38	40.6	1.0	1356	4	US-09-252-991A-9233	Sequence 9233, Ap
c	39	40.6	1.0	2715	4	US-09-252-991A-8715	Sequence 8715, Ap
	40	40.6	1.0	3126	4	US-09-252-991A-9066	Sequence 9066, Ap
	41	40.6	1.0	4143	4	US-09-252-991A-9151	Sequence 9151, Ap
c	42	40.6	1.0	4212	4	US-09-252-991A-8929	Sequence 8929, Ap
	43	40.4	1.0	3572	4	US-09-575-574-3	Sequence 3, Appli
	44	40.2	1.0	1026	3	US-07-751-891B-24	Sequence 24, Appl
	45	40.2	1.0	5222	3	US-07-751-891B-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1
US-09-245-041-14
; Sequence 14, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 8589
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-245-041-14

Query Match 98.0%; Score 3988.4; DB 3; Length 8589;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 4013; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```
Qy      15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74
           || |||    |  ||   | ||  |  |  | | | |   |||| ||   || || ||
Db     234 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAA 293

Qy      75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134
           |    | |  | ||  ||||||||||||||||||||||||||||||||||||||||||||
Db     294 GGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 353

Qy     135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     354 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 413

Qy     195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     414 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 473

Qy     255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     474 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 533

Qy     315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        534 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 593
Qy        375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434
Db        594 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 653
Qy        435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494
Db        654 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 713
Qy        495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554
Db        714 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 773
Qy        555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614
Db        774 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 833
Qy        615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 674
Db        834 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 893
Qy        675 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 734
Db        894 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 953
Qy        735 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 794
Db        954 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 1013
Qy        795 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 854
Db       1014 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 1073
Qy        855 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 914
Db       1074 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 1133
Qy        915 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 974
Db       1134 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 1193
Qy        975 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1034
Db       1194 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1253
Qy       1035 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1094
Db       1254 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1313
Qy       1095 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1154
Db       1314 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1373
Qy       1155 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1214
Db       1374 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1433
Qy       1215 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1274
Db       1434 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1493
Qy       1275 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1334
Db       1494 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1553
Qy       1335 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1394
Db       1554 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1613
Qy       1395 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1454
Db       1614 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1673

Qy       1455 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1514
Db       1674 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1733
Qy       1515 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1574
Db       1734 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1793
Qy       1575 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGC 1634
Db       1794 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGC 1853
Qy       1635 AGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGA 1694
Db       1854 AGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGA 1913
Qy       1695 CATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGC 1754
Db       1914 CATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGC 1973
Qy       1755 AGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGC 1814
Db       1974 AGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGC 2033
Qy       1815 GCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCT 1874
Db       2034 GCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCT 2093
Qy       1875 TGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGA 1934
Db       2094 TGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGA 2153
Qy       1935 CTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGAT 1994
Db       2154 CTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGAT 2213
Qy       1995 CTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAA 2054
Db       2214 CTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAA 2273
Qy       2055 GACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGA 2114
Db       2274 GACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGA 2333
Qy       2115 GTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATG 2174
Db       2334 GTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATG 2393
Qy       2175 TTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCA 2234
Db       2394 TTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCA 2453
Qy       2235 CAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCT 2294
Db       2454 CAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCT 2513
Qy       2295 GCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCT 2354
Db       2514 GCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCT 2573
Qy       2355 TCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTT 2414
Db       2574 TCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTT 2633
Qy       2415 ACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACC 2474
Db       2634 ACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACC 2693
Qy       2475 CAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGA 2534
Db       2694 CAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGA 2753
```

```
Qy       2535 AAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATG 2594
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2754 AAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATG 2813

Qy       2595 TGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGT 2654
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2814 TGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGT 2873

Qy       2655 GGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGAT 2714
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2874 GGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGAT 2933

Qy       2715 GAGCACCTGCCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCA 2774
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2934 GAGCACCTGCCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCA 2993

Qy       2775 ACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGG 2834
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2994 ACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGG 3053

Qy       2835 TTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACA 2894
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3054 TTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACA 3113

Qy       2895 GCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCA 2954
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3114 GCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCA 3173

Qy       2955 CTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAA 3014
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3174 CTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAA 3233

Qy       3015 GTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGA 3074
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3234 GTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGA 3293

Qy       3075 TCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAA 3134
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3294 TCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAA 3353

Qy       3135 CACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCT 3194
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3354 CACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCT 3413

Qy       3195 ATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCT 3254
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3414 ATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCT 3473

Qy       3255 TCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGC 3314
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3474 TCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGC 3533

Qy       3315 TATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGC 3374
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3534 TATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGC 3593

Qy       3375 CTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGC 3434
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3594 CTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGC 3653

Qy       3435 TGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTC 3494
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3654 TGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTC 3713

Qy       3495 TAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAA 3554
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3714 TAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAA 3773

Qy       3555 TTTCACCTGGCCCATCAAAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCT 3614
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3774 TTTCACCTGGCCCATCAAAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCT 3833

Qy       3615 GGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGT 3674
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3834 GGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGT 3893

Qy       3675 GGTTTGGAAGATCAAACAAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGA 3734
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3894 GGTTTGGAAGATCAAACAAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGA 3953

Qy       3735 GATGCAACAGATGGCCAGCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGA 3794
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3954 GATGCAACAGATGGCCAGCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGA 4013

Qy       3795 GGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGA 3854
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4014 GGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGA 4073

Qy       3855 GCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCT 3914
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4074 GCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCT 4133

Qy       3915 GGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACAT 3974
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4134 GGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACAT 4193

Qy       3975 TTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCA 4034
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4194 TTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCA 4253

Qy       4035 GCAGCCCCCTGCACAGCCTGGGACCTGCATCTGA 4068
             ||||||||||||||||||||||||||||||||||
Db       4254 GCAGCCCCCTGCACAGCCTGGGACCTGCATCTGA 4287
```

RESULT 2
US-09-245-041-16
; Sequence 16, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 4072
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-245-041-16

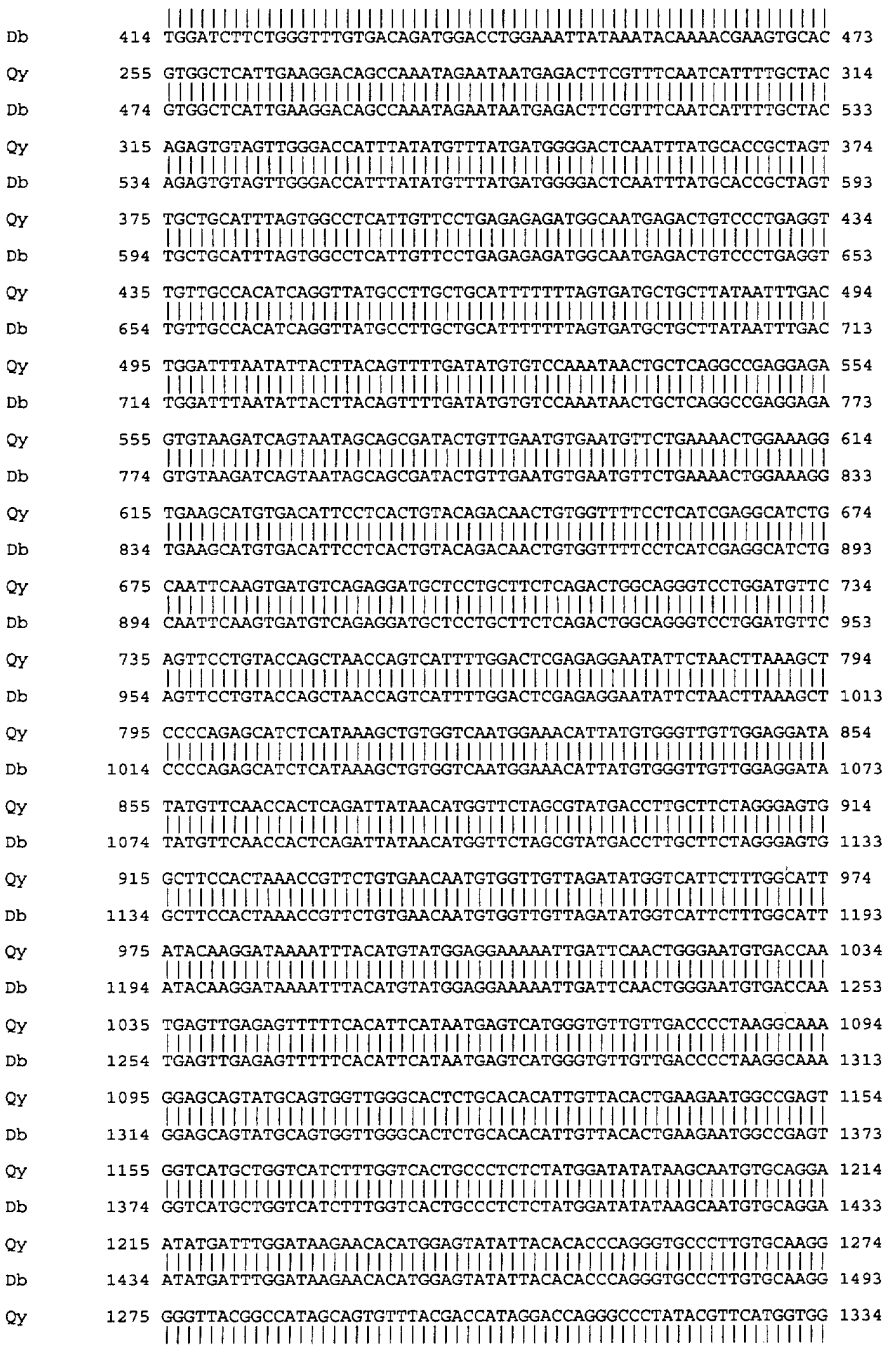
Query Match 86.0%; Score 3499.8; DB 3; Length 4072;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 3525; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

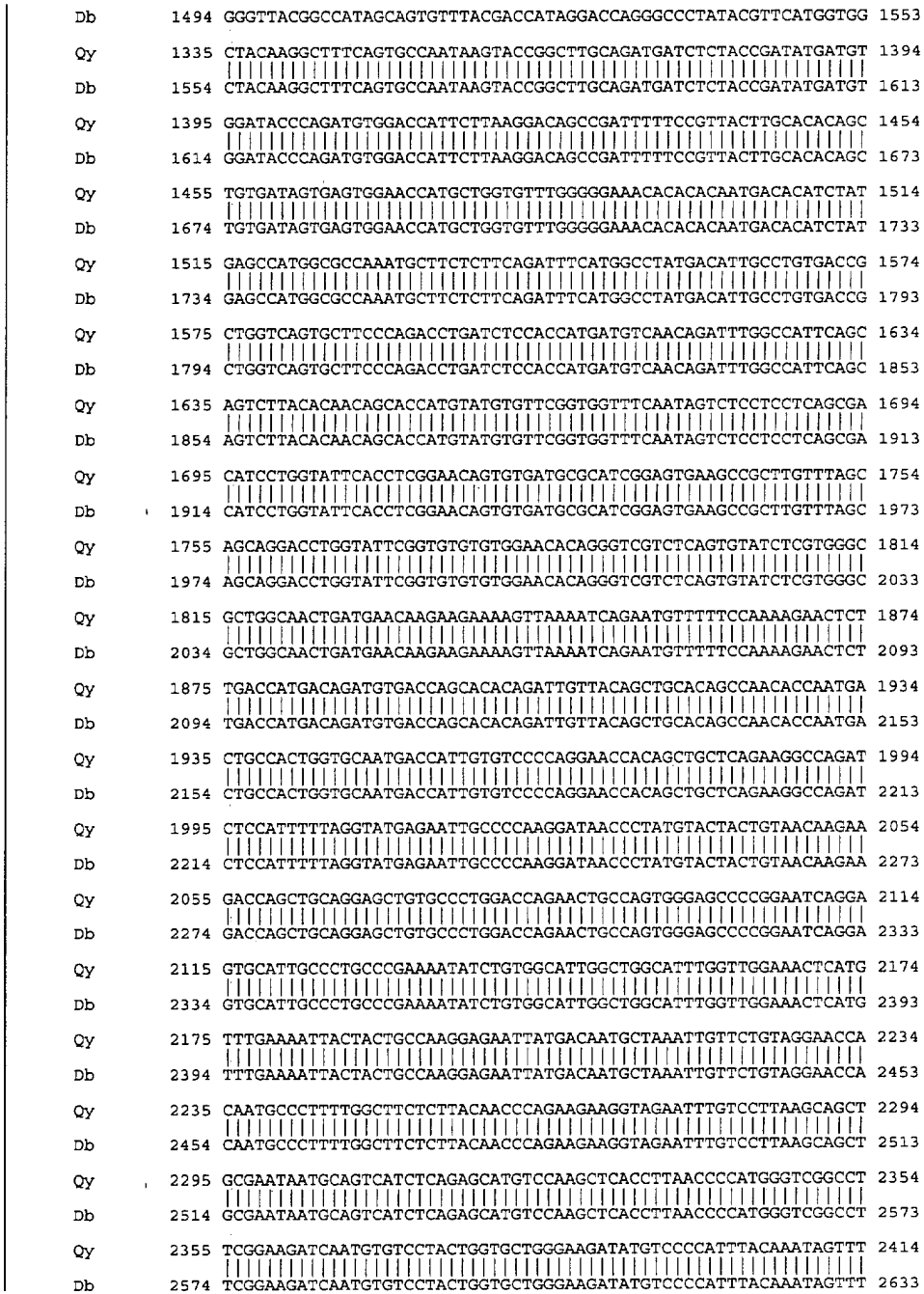
```
Qy         15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74
              || |||    |  ||   | ||  |  |  |  | | |  |||| ||   || || ||
Db        234 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAA 293

Qy         75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134
              |   | |  | ||  ||||||||||||||||||||||||||||||||||||||||||||||
Db        294 GGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 353

Qy        135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        354 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 413

Qy        195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254
```


Db 414 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 473
Qy 255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314
Db 474 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 533
Qy 315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374
Db 534 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 593
Qy 375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434
Db 594 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 653
Qy 435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494
Db 654 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 713
Qy 495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554
Db 714 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 773
Qy 555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614
Db 774 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 833
Qy 615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 674
Db 834 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 893
Qy 675 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 734
Db 894 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 953
Qy 735 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 794
Db 954 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 1013
Qy 795 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 854
Db 1014 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 1073
Qy 855 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 914
Db 1074 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 1133
Qy 915 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 974
Db 1134 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 1193
Qy 975 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1034
Db 1194 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1253
Qy 1035 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1094
Db 1254 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1313
Qy 1095 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1154
Db 1314 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1373
Qy 1155 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1214
Db 1374 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1433
Qy 1215 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1274
Db 1434 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1493
Qy 1275 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1334


Db 1494 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1553
Qy 1335 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1394
Db 1554 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1613
Qy 1395 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1454
Db 1614 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1673
Qy 1455 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1514
Db 1674 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1733
Qy 1515 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1574
Db 1734 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1793
Qy 1575 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGC 1634
Db 1794 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGC 1853
Qy 1635 AGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGA 1694
Db 1854 AGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGA 1913
Qy 1695 CATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGC 1754
Db 1914 CATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGC 1973
Qy 1755 AGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGC 1814
Db 1974 AGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGC 2033
Qy 1815 GCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCT 1874
Db 2034 GCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCT 2093
Qy 1875 TGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGA 1934
Db 2094 TGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGA 2153
Qy 1935 CTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGAT 1994
Db 2154 CTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGAT 2213
Qy 1995 CTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAA 2054
Db 2214 CTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAA 2273
Qy 2055 GACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGA 2114
Db 2274 GACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGA 2333
Qy 2115 GTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATG 2174
Db 2334 GTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATG 2393
Qy 2175 TTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCA 2234
Db 2394 TTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCA 2453
Qy 2235 CAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCT 2294
Db 2454 CAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCT 2513
Qy 2295 GCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCT 2354
Db 2514 GCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCT 2573
Qy 2355 TCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTT 2414
Db 2574 TCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTT 2633

```
Qy       2415 ACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACC 2474
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2634 ACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACC 2693

Qy       2475 CAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGA 2534
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2694 CAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGA 2753

Qy       2535 AAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATG 2594
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2754 AAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATG 2813

Qy       2595 TGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGT 2654
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2814 TGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGT 2873

Qy       2655 GGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGAT 2714
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2874 GGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGAT 2933

Qy       2715 GAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCA 2774
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2934 GAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCA 2993

Qy       2775 ACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGG 2834
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2994 ACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGG 3053

Qy       2835 TTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACA 2894
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3054 TTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACA 3113

Qy       2895 GCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCA 2954
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3114 GCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCA 3173

Qy       2955 CTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAA 3014
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3174 CTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAA 3233

Qy       3015 GTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGA 3074
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3234 GTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGA 3293

Qy       3075 TCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAA 3134
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3294 TCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAA 3353

Qy       3135 CACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCT 3194
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3354 CACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCT 3413

Qy       3195 ATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCT 3254
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3414 ATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCT 3473

Qy       3255 TCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGC 3314
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3474 TCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGC 3533

Qy       3315 TATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGC 3374
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3534 TATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGC 3593

Qy       3375 CTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGC 3434
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3594 CTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGC 3653

Qy       3435 TGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTC 3494
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3654 TGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTC 3713

Qy       3495 TAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAA 3554
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3714 TAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAA 3773

Qy       3555 TTTCACCTGGCCCATCAAAATTCAGAT 3581
             |||||||||||||||||||||||||||
Db       3774 TTTCACCTGGCCCATCAAAATTCAGGT 3800
```

RESULT 3
US-09-245-041-1
; Sequence 1, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 8827
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-1

Query Match 81.8%; Score 3326.4; DB 3; Length 8827;
Best Local Similarity 91.1%; Pred. No. 0;
Matches 3534; Conservative 0; Mismatches 346; Indels 0; Gaps 0;

```
Qy        189 ACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAA 248
              |||||||| |||||||| ||||| |||||||||||||| |||||||||| || |||||
Db        677 ACTAACTGGCTCTTCTGGATTTGTAACAGATGGACCTGGGAATTATAAATATAAGACGAA 736

Qy        249 GTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTT 308
              |||||| |||||||||||||||||||||||||||||||||||||||||| ||||| |||||
Db        737 GTGCACATGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGCTTCAACCATTT 796

Qy        309 TGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACC 368
              |||||||| ||||| |||||||||||||||||||||||||||||||||| || ||||||
Db        797 TGCTACAGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATCTACGCACC 856

Qy        369 GCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCC 428
              ||  ||||||| |||||||||||||||||||||||| |||||||||||||||| | ||
Db        857 TCTGATTGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCTCC 916

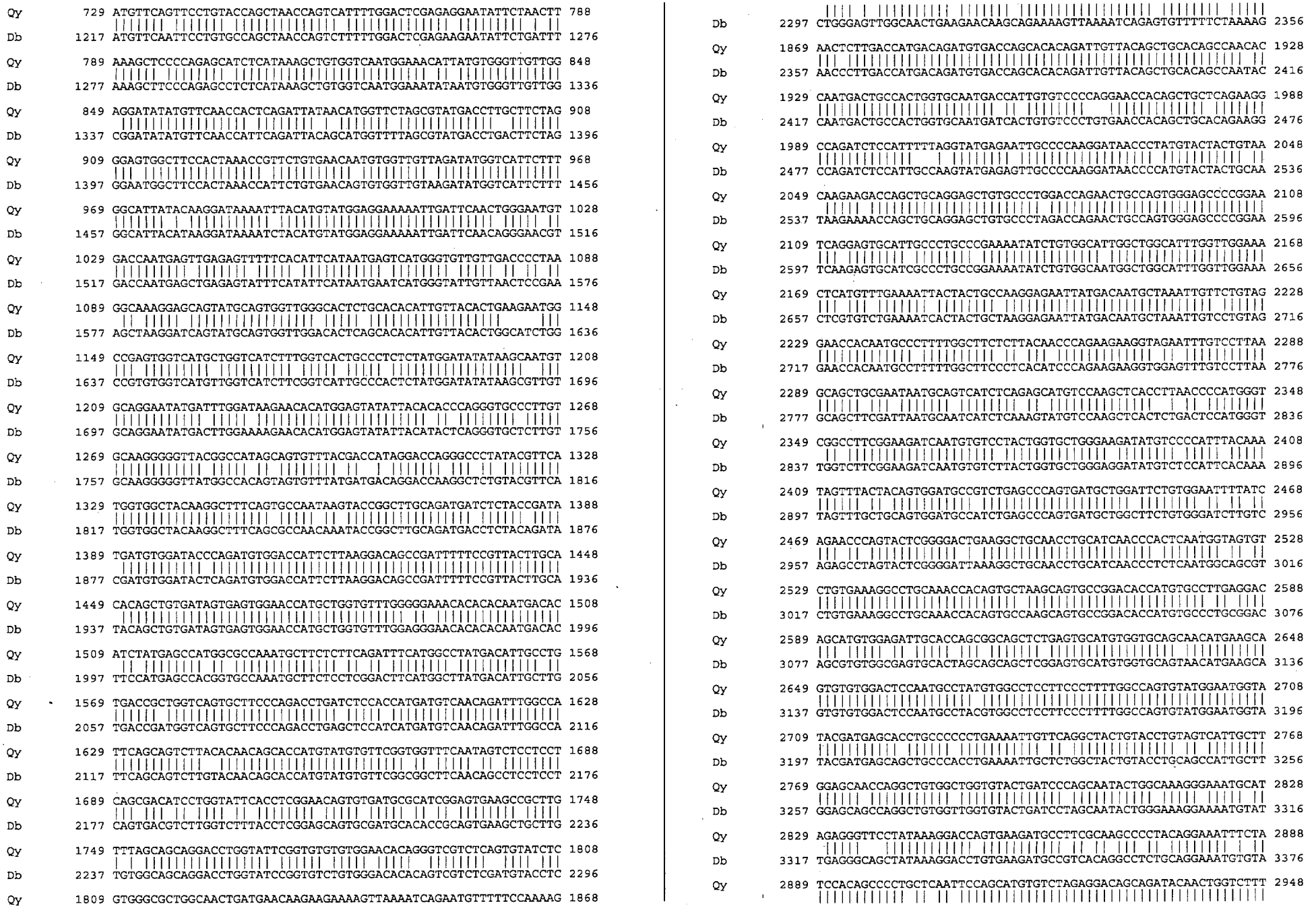
Qy        429 TGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAA 488
              ||||||   || ||| |||||||||| |||||||||||| |||||||||||||||||||
Db        917 TGAGGTCACTGTCACTTCAGGTTATGCACTGCTGCATTTTTTCAGTGATGCTGCTTATAA 976

Qy        489 TTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCG 548
              | |||||||||||||| ||||||| |||||| |||||||| ||||| ||||||||||
Db        977 TCTGACTGGATTTAATATCACTTACAATTTTGACATGTGTCCGAATAATTGCTCAGGCCG 1036

Qy        549 AGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTG 608
              |||||||||||||| |||||| |||||    ||||||| |||||||||||||||||||
Db        1037 AGGAGAGTGTAAGAGCAGTAACAGCAGCAGCGCTGTTGAGTGTGAATGTTCTGAAAACTG 1096

Qy        609 GAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGG 668
              |||||| ||  | |||||||||||||||||||||||||||||||||| |||||| |||||
Db        1097 GAAAGGGGAGTCGTGTGACATTCCTCACTGTACAGACAACTGTGGCTTTCCTCACCGAGG 1156

Qy        669 CATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGG 728
              |||||| ||| |||| ||| |||||| |||||||||| | |||||||||||||||||
Db        1157 CATCTGTAATGCAAGCGATACCAGAGGGTGCTCCTGCTTTCCTCACTGGCAGGGTCCTGG 1216
```



```
Db       3377 TCCACAGCCCCTTCTGAACTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTT 3436
Qy       2949 CATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTG 3008
Db       3437 CATTCACTGTCCAGCTTGCCAGTGCAACGGACACAGCAAATGCATCAACCAGAGTATCTG 3496
Qy       3009 TGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTA 3068
Db       3497 TGAGAAGTGTGAGGACCTGACCACGGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTA 3556
Qy       3069 CGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCT 3128
Db       3557 TGGTGACCCGACTAATGGAGGCAAATGTCAGCCATGCAAGTGCAATGGGCACGCATCACT 3616
Qy       3129 GTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTG 3188
Db       3617 GTGCAACACCAACACCGGCAAGTGCTTCTGTACCACCAAAGGTGTCAAGGGGGACGAGTG 3676
Qy       3189 CCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTA 3248
Db       3677 CCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAAAGGAACATGCTACTA 3736
Qy       3249 TACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTA 3308
Db       3737 TACCCTTCTCATTGACTATCAGTTCACCTTTAGCCTGTCCCAGGAAGACGACCGCTACTA 3796
Qy       3309 CACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCAT 3368
Db       3797 CACAGCCATCAACTTTGTGGCTACTCCTGATGAACAAAACAGGGATTTGGACATGTTCAT 3856
Qy       3369 CAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAAC 3428
Db       3857 CAATGCCTCCAAAAACTTCAACCTCAACATCACCTGGGCCACCAGCTTCCCAGCCGGAAC 3916
Qy       3429 CCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAG 3488
Db       3917 CCAGACTGGAGAAGAGGTGCCTGTTGTTTCAAAAACCAACATCAAGGAATACAAAGATAG 3976
Qy       3489 TTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGT 3548
Db       3977 CTTCTCTAATGAGAAATTTGATTTTCGCAACCATCCAAACATCACTTTCTTTGTTTATGT 4036
Qy       3549 CAGTAATTTCACCTGGCCCATCAAAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTAT 3608
Db       4037 CAGTAATTTCACTTGGCCCATCAAAATTCAGATTGCCTTCTCCCAGCACAGCAACTTCAT 4096
Qy       3609 GGACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGC 3668
Db       4097 GGACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCGCTGCTTCTGGTGGC 4156
Qy       3669 TGCTGTGGTTTGGAAGATCAAACAAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCT 3728
Db       4157 TGCAGTGGTCTGGAAGATCAAGCAGAGCTGTTGGGCATCCAGGCGGAGAGAGCAACTTCT 4216
Qy       3729 TCGAGAGATGCAACAGATGGCCAGCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAAC 3788
Db       4217 TCGGGAGATGCAACAGATGGCCAGCCGCCCCTTTGCTTCTGTAAACGTTGCCTTGGAAAC 4276
Qy       3789 AGATGAGGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGC 3848
Db       4277 AGATGAGGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGC 4336
Qy       3849 ACTGGAGCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCG 3908
Db       4337 ACTGGAGCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCG 4396
Qy       3909 AGGCCTGGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGT 3968
Db       4397 AGGCCTGGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGT 4456
Qy       3969 GGACATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCG 4028
Db       4457 GGACATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCG 4516
```

```
Qy       4029 GAAGCAGCAGCCCCCTGCACAGCCTGGGACCTGCATCTGA 4068
Db       4517 GAAGCAGCAGCCCCCTGCACAGCCTGGGACCTGCATCTGA 4556
```

RESULT 4
US-09-245-041-18
; Sequence 18, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 2625
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-245-041-18

Query Match 50.4%; Score 2051.8; DB 3; Length 2625;
Best Local Similarity 97.1%; Pred. No. 0;
Matches 2089; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

```
Qy         15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74
Db        234 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAA 293
Qy         75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134
Db        294 GGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 353
Qy        135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194
Db        354 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 413
Qy        195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254
Db        414 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 473
Qy        255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314
Db        474 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 533
Qy        315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374
Db        534 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 593
Qy        375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434
Db        594 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 653
Qy        435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494
Db        654 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 713
Qy        495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554
Db        714 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 773
Qy        555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614
Db        774 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 833
```

```
Qy       615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 674
Db       834 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 893

Qy       675 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 734
Db       894 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 953

Qy       735 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 794
Db       954 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCT 1013

Qy       795 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 854
Db      1014 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATA 1073

Qy       855 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 914
Db      1074 TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 1133

Qy       915 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 974
Db      1134 GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATT 1193

Qy       975 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1034
Db      1194 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAA 1253

Qy      1035 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1094
Db      1254 TGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAA 1313

Qy      1095 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1154
Db      1314 GGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGT 1373

Qy      1155 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1214
Db      1374 GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1433

Qy      1215 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1274
Db      1434 ATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1493

Qy      1275 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1334
Db      1494 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGG 1553

Qy      1335 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1394
Db      1554 CTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1613

Qy      1395 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1454
Db      1614 GGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGC 1673

Qy      1455 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1514
Db      1674 TGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTAT 1733

Qy      1515 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1574
Db      1734 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCG 1793

Qy      1575 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGC 1634
Db      1794 CTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGC 1853

Qy      1635 AGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGA 1694
Db      1854 AGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGA 1913
```

```
Qy      1695 CATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGC 1754
Db      1914 CATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGC 1973

Qy      1755 AGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGC 1814
Db      1974 AGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGC 2033

Qy      1815 GCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCT 1874
Db      2034 GCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCT 2093

Qy      1875 TGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGA 1934
Db      2094 TGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGA 2153

Qy      1935 CTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGAT 1994
Db      2154 CTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGAT 2213

Qy      1995 CTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAA 2054
Db      2214 CTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAA 2273

Qy      2055 GACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGA 2114
Db      2274 GACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGA 2333

Qy      2115 GTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGG 2165
Db      2334 GTGCATTGCCCTGCCCGGTAGGCCTTGCAGGGTCATCTTGGTGTGTGTGGG 2384
```

RESULT 5
US-09-245-041-12
; Sequence 12, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 6370
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-245-041-12

Query Match 45.6%; Score 1853; DB 3; Length 6370;
Best Local Similarity 96.3%; Pred. No. 0;
Matches 1980; Conservative 7; Mismatches 47; Indels 23; Gaps 8;

```
Qy      2035 ATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAG 2094
Db         1 ATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAG 60

Qy      2095 TGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGG 2154
Db        61 TGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGG 120

Qy      2155 CATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCT 2214
Db       121 CATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCT 180
```

```
Qy       2215 AAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTA 2274
Db        181 AAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTA 240

Qy       2275 GAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACC 2334
Db        241 GAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACC 300

Qy       2335 TTAACCCCATGGGTCGGCCTT---CGGAAGATCAATGTGTCCTACTGG-----TGCTGGG 2386
Db        301 TTAACCCCATGGGTCGGGCCTTCGGGAAGGTYCAATGTGTCCTACTKGGTGCTGGGGAAG 360

Qy       2387 AAGATATGTCCCCATTTACAAATAG-TTTACTACAGTGGATGCC----GTCTGAGCCCAG 2441
Db        361 GATATGKTCCCCATTTTACAAATAGTTTTACTACAGTGGGATGSCCGTCTTGAGGCCCAG 420

Qy       2442 TGATGCT--GGATTCTGTGGAATTTT-----ATCAGAACCCAGTACTCGGGGACTGAAGG 2494
Db        421 TGTTGCTTGGRATTCTGTGGGAATTTTATTCAGGAACCCAGTTACTTCGGGGACTGAAGG 480

Qy       2495 CTGCAACCTGCAT--CAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAG 2552
Db        481 CTGCAACCTGCATTCAACCCACTYMAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAG 540

Qy       2553 TGCTAA-GCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCG 2611
Db        541 TGCTAAGGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCG 600

Qy       2612 GCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATG 2671
Db        601 GCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATG 660

Qy       2672 TGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTG 2731
Db        661 TGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTG 720

Qy       2732 AAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGT 2791
Db        721 AAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGT 780

Qy       2792 GTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAG 2851
Db        781 GTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAG 840

Qy       2852 TGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCA 2911
Db        841 TGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCA 900

Qy       2912 GCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAAT 2971
Db        901 GCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAAT 960

Qy       2972 GCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCA 3031
Db        961 GCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCA 1020

Qy       3032 CAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGA 3091
Db       1021 CAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGA 1080

Qy       3092 AATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGT 3151
Db       1081 AATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGT 1140

Qy       3152 GCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATC 3211
Db       1141 GCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATC 1200

Qy       3212 GATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGT 3271
Db       1201 GATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGT 1260

Qy       3272 TCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTA 3331
Db       1261 TCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTA 1320

Qy       3332 CTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACC 3391
Db       1321 CTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACC 1380

Qy       3392 TCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTG 3451
Db       1381 TCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTG 1440

Qy       3452 TTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATT 3511
Db       1441 TTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATT 1500

Qy       3512 TTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCA 3571
Db       1501 TTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCA 1560

Qy       3572 AAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCTGGTACAGTTCTTCGTGA 3631
Db       1561 AAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCTGGTACAGTTCTTCGTGA 1620

Qy       3632 CTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGTGGTTTGGAAGATCAAAC 3691
Db       1621 CTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGTGGTTTGGAAGATCAAAC 1680

Qy       3692 AAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGAGATGCAACAGATGGCCA 3751
Db       1681 AAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGAGATGCAACAGATGGCCA 1740

Qy       3752 GCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTA 3811
Db       1741 GCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTA 1800

Qy       3812 TTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGAGCCGTGTTTTGGCAACA 3871
Db       1801 TTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGAGCCGTGTTTTGGCAACA 1860

Qy       3872 AAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCTGGGTGGCATCCCTCCTC 3931
Db       1861 AAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCTGGGTGGCATCCCTCCTC 1920

Qy       3932 CTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGA 3991
Db       1921 CTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGA 1980

Qy       3992 TAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGC 4051
Db       1981 TAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGC 2040

Qy       4052 CTGGGACCTGCATCTGA 4068
Db       2041 CTGGGACCTGCATCTGA 2057
```

RESULT 6
US-09-245-041-8
; Sequence 8, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131

```
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
;   LENGTH: 2419
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-245-041-8

  Query Match            35.7%;  Score 1451.8;  DB 3;  Length 2419;
  Best Local Similarity  89.6%;  Pred. No. 0;
  Matches 1561;  Conservative    0;  Mismatches  182;  Indels    0;  Gaps    0;

Qy       189 ACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAA 248
Db       677 ACTAACTGGCTCTTCTGGATTTGTAACAGATGGACCTGGGAATTATAAATATAAGACGAA 736

Qy       249 GTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTT 308
Db       737 GTGCACATGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGCTTCAACCATTT 796

Qy       309 TGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACC 368
Db       797 TGCTACAGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATCTACGCACC 856

Qy       369 GCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCC 428
Db       857 TCTGATTGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCTCC 916

Qy       429 TGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAA 488
Db       917 TGAGGTCACTGTCACTTCAGGTTATGCACTGCTGCATTTTTTCAGTGATGCTGCTTATAA 976

Qy       489 TTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCG 548
Db       977 TCTGACTGGATTTAATATCACTTACAATTTTGACATGTGTCCGAATAATTGCTCAGGCCG 1036

Qy       549 AGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTG 608
Db      1037 AGGAGAGTGTAAGAGCAGTAACAGCAGCAGCGCTGTTGAGTGTGAATGTTCTGAAAACTG 1096

Qy       609 GAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGG 668
Db      1097 GAAAGGGGAGTCGTGTGACATTCCTCACTGTACAGACAACTGTGGCTTTCCTCACCGAGG 1156

Qy       669 CATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGG 728
Db      1157 CATCTGTAATGCAAGCGATACCAGAGGGTGCTCCTGCTTTCCTCACTGGCAGGGTCCTGG 1216

Qy       729 ATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTT 788
Db      1217 ATGTTCAATTCCTGTGCCAGCTAACCAGTCTTTTTGGACTCGAGAAGAATATTCTGATTT 1276

Qy       789 AAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGG 848
Db      1277 AAAGCTTCCCAGAGCCTCTCATAAAGCTGTGGTCAATGGAAATATAATGTGGGTTGTTGG 1336

Qy       849 AGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAG 908
Db      1337 CGGATATATGTTCAACCATTCAGATTACAGCATGGTTTTAGCGTATGACCTGACTTCTAG 1396

Qy       909 GGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTT 968
Db      1397 GGAATGGCTTCCACTAAACCATTCTGTGAACAGTGTGGTTGTAAGATATGGTCATTCTTT 1456

Qy       969 GGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGT 1028
Db      1457 GGCATTACATAAGGATAAAATCTACATGTATGGAGGAAAAATTGATTCAACAGGGAACGT 1516

Qy      1029 GACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAA 1088
Db      1517 GACCAATGAGCTGAGAGTATTTCATATTCATAATGAATCATGGGTATTGTTAACTCCGAA 1576

Qy      1089 GGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGG 1148
Db      1577 AGCTAAGGATCAGTATGCAGTGGTTGGACACTCAGCACACATTGTTACACTGGCATCTGG 1636

Qy      1149 CCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGT 1208
Db      1637 CCGTGTGGTCATGTTGGTCATCTTCGGTCATTGCCCACTCTATGGATATATAAGCGTTGT 1696

Qy      1209 GCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGT 1268
Db      1697 GCAGGAATATGACTTGGAAAAGAACACATGGAGTATATTACATACTCAGGGTGCTCTTGT 1756

Qy      1269 GCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCA 1328
Db      1757 GCAAGGGGGTTATGGCCACAGTAGTGTTTATGATGACAGGACCAAGGCTCTGTACGTTCA 1816

Qy      1329 TGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATA 1388
Db      1817 TGGTGGCTACAAGGCTTTCAGCGCCAACAAATACCGGCTTGCAGATGACCTCTACAGATA 1876

Qy      1389 TGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCA 1448
Db      1877 CGATGTGGATACTCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCA 1936

Qy      1449 CACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACAC 1508
Db      1937 TACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGAGGGAACACACACAATGACAC 1996

Qy      1509 ATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTG 1568
Db      1997 TTCCATGAGCCACGGTGCCAAATGCTTCTCCTCGGACTTCATGGCTTATGACATTGCTTG 2056

Qy      1569 TGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCA 1628
Db      2057 TGACCGATGGTCAGTGCTTCCCAGACCTGAGCTCCATCATGATGTCAACAGATTTGGCCA 2116

Qy      1629 TTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCT 1688
Db      2117 TTCAGCAGTCTTGTACAACAGCACCATGTATGTGTTCGGCGGCTTCAACAGCCTCCTCCT 2176

Qy      1689 CAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTG 1748
Db      2177 CAGTGACGTCTTGGTCTTTACCTCGGAGCAGTGCGATGCACACCGCAGTGAAGCTGCTTG 2236

Qy      1749 TTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTC 1808
Db      2237 TGTGGCAGCAGGACCTGGTATCCGGTGTCTGTGGGACACACAGTCGTCTCGATGTACCTC 2296

Qy      1809 GTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAG 1868
Db      2297 CTGGGAGTTGGCAACTGAAGAACAAGCAGAAAAGTTAAAATCAGAGTGTTTTTCTAAAAG 2356

Qy      1869 AACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACAC 1928
Db      2357 AACCCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAATAC 2416

Qy      1929 CAA 1931
Db      2417 CAA 2419

RESULT 7
US-09-245-041-10
; Sequence 10, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT: Nagle, D.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
;  TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
;  FILE REFERENCE: 7853-136
;  CURRENT APPLICATION NUMBER: US/09/245,041
;  CURRENT FILING DATE: 1999-02-05
;  EARLIER APPLICATION NUMBER: 60/093,630
;  EARLIER FILING DATE: 1998-07-21
```

; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 1051
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-10

Query Match 11.4%; Score 463; DB 3; Length 1051;
Best Local Similarity 85.2%; Pred. No. 4e-127;
Matches 517; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

```
Qy        15 GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCGCTCGCGGGCAG 74
             || |||    |  ||   | ||  |  |  | | | |||| ||   || || ||
Db       441 GGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAA 500

Qy        75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGT 134
             |    | |  | ||  ||||||||||||||| ||||||||||||||||||||||||||||
Db       501 GGAATGTGACCGGCCGTGTGTCAACGGCGGCCGCTGCAACCCTGGCACCGGCCAGTGCGT 560

Qy       135 CTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 194
             ||||||| | |||||||||||||||||||||||||||||||||||||||||||||||||
Db       561 CTGCCCCACGGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAAC 620

Qy       195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCAC 254
             ||| |||||||| ||||| |||||||||||||| |||||||||| || ||||||||||
Db       621 TGGCTCTTCTGGATTTGTAACAGATGGACCTGGGAATTATAAATATAAGACGAAGTGCAC 680

Qy       255 GTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTAC 314
              |||||||||||||||||||||||||||||||||||||||| ||||| ||||||||||
Db       681 ATGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGCTTCAACCATTTTGCTAC 740

Qy       315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT 374
             ||| ||||| |||||||||||||||||||||||||||||||| || ||||| ||   |
Db       741 AGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATCTACGCACCTCTGAT 800

Qy       375 TGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434
             |||||| |||||||||||||||||||||||| |||||||||||||| | ||||||||
Db       801 TGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCTCCTGAGGT 860

Qy       435 TGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGAC 494
                || ||| |||||||||||  |||||||||||| |||||||||||||||||| ||||
Db       861 CACTGTCACTTCAGGTTATGCACTGCTGCATTTTTTCAGTGATGCTGCTTATAATCTGAC 920

Qy       495 TGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGA 554
             |||||||||||| ||||||| |||||| |||||||| ||||| ||||||||||||||||
Db       921 TGGATTTAATATCACTTACAATTTTGACATGTGTCCGAATAATTGCTCAGGCCGAGGAGA 980

Qy       555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 614
             |||||||| |||||| ||||||    ||||||| ||||||||||||||||||||||||
Db       981 GTGTAAGAGCAGTAACAGCAGCAGCGCTGTTGAGTGTGAATGTTCTGAAAACTGGAAAGG 1040

Qy       615 TGAAGCA 621
                 | | |
Db      1041 GGCCGGA 1047
```

RESULT 8
US-09-702-705-966/c
; Sequence 966, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-966

Query Match 6.0%; Score 246; DB 4; Length 246;
Best Local Similarity 100.0%; Pred. No. 5.2e-63;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      2293 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 2352
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy      2353 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 2412
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy      2413 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 2472
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67

Qy      2473 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 2532
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy      2533 GAAAGG 2538
             ||||||
Db          6 GAAAGG 1
```

RESULT 9
US-09-736-457-966/c
; Sequence 966, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-966

Query Match 6.0%; Score 246; DB 4; Length 246;
Best Local Similarity 100.0%; Pred. No. 5.2e-63;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      2293 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 2352
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187
```

```
Qy       2353 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 2412
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy       2413 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 2472
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67

Qy       2473 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 2532
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy       2533 GAAAGG 2538
             ||||||
Db           6 GAAAGG 1


RESULT 10
US-09-614-124B-966/c
; Sequence 966, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
;  APPLICANT: Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.478C9
;  CURRENT APPLICATION NUMBER: US/09/614,124B
;  CURRENT FILING DATE:  2001-07-11
;  NUMBER OF SEQ ID NOS: 1668
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 966
;   LENGTH: 246
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-614-124B-966

  Query Match             6.0%;  Score 246;  DB 4;  Length 246;
  Best Local Similarity   100.0%;  Pred. No. 5.2e-63;
  Matches  246;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2293 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 2352
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy       2353 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 2412
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy       2413 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 2472
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67

Qy       2473 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 2532
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy       2533 GAAAGG 2538
             ||||||
Db           6 GAAAGG 1


RESULT 11
US-09-671-325-966/c
; Sequence 966, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Fan, Liqun
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.478C12
;  CURRENT APPLICATION NUMBER: US/09/671,325
;  CURRENT FILING DATE:  2000-09-26
;  NUMBER OF SEQ ID NOS:  1825
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 966
;   LENGTH: 246
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-671-325-966

  Query Match             6.0%;  Score 246;  DB 4;  Length 246;
  Best Local Similarity   100.0%;  Pred. No. 5.2e-63;
  Matches  246;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2293 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 2352
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy       2353 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 2412
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy       2413 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 2472
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67

Qy       2473 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 2532
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy       2533 GAAAGG 2538
             ||||||
Db           6 GAAAGG 1


RESULT 12
US-09-833-381-1916
; Sequence 1916, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
;  APPLICANT: Robison, Keith E.
;  TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
;  FILE REFERENCE: 5800-119
;  CURRENT APPLICATION NUMBER: US/09/833,381
;  CURRENT FILING DATE:  2001-04-11
;  PRIOR APPLICATION NUMBER: 09/516,448
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 2050
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1916
;   LENGTH: 415
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(415)
;   OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1916

  Query Match             5.2%;  Score 212.8;  DB 4;  Length 415;
```

Best Local Similarity 73.6%; Pred. No. 5.8e-53;
Matches 271; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

```
Qy        200 CTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGGC 259
              | || || | | | |||||||| ||    || |||||||| ||||| || || || ||||
Db          5 CGTCCGGATATTTAACAGATGGCCCAATTAACTATAAATATAAAACTAAATGTACTTGGC  64

Qy        260 TCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAGT 319
              |||||||||  | ||||||  | |  | ||| |  | ||||||||||||||||||| |
Db         65 TCATTGAAGGCTATCCAAATGCAGTGTTAAGATTAAGATTCAATCATTTTGCTACAGAAT 124

Qy        320 GTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCTG 379
              |||| ||||| ||| | ||||||||||||||| || ||||| ||||||||  || | ||||
Db        125 GTAGCTGGGATCATATGTATGTTTATGATGGAGATTCAATATATGCACCTTTAATAGCTG 184

Qy        380 CATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTG 439
              | |||||||  | || || ||||| | |   |||||||| ||||| |||||| ||||||
Db        185 TACTTAGTGGTTTGATAGTCCCTGAAATAAGGGGCAATGAAACTGTGCCTGAAGTTGTTA 244

Qy        440 CCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGAT 499
              | ||||| || |||||  || | |||||||||||||||||||| |||||| | ||||| |
Db        245 CTACATCTGGCTATGCACTGTTACATTTTTTTAGTGATGCTGCGTATAATCTAACTGGTT 304

Qy        500 TTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGTA 559
              | || |||   ||     | ||   ||||| || || ||||| || |  ||  ||||||
Db        305 TCAACATTTTCTATTCAATCAATTCTTGTCCTAACAATTGCTCTGGTCATGGGAAGTGTA 364

Qy        560 AGATCAGT 567
              |  |||
Db        365 CAACTAGT 372
```

RESULT 13
US-09-245-041-5
; Sequence 5, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 90050
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-5

Query Match 5.0%; Score 202.8; DB 3; Length 90050;
Best Local Similarity 81.8%; Pred. No. 2.8e-48;
Matches 234; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

```
Qy        947 TTGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAA 1006
              | |||  | |    | | |    | || |  |||||||||| ||||||||||||||||
Db       6931 TGGTTTTAAAAATCCTGTTTACATCTTTCTTATAGGATAAAATCTACATGTATGGAGGAA 6990

Qy       1007 AAATTGATTCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGT 1066
              |||||||||||| |||| |||||||||||| ||||||| ||||| ||||||||||| |
Db       6991 AAATTGATTCAACAGGGAACGTGACCAATGAGCTGAGAGTATTTCATATTCATAATGAAT 7050

Qy       1067 CATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCAC 1126
              ||||||| ||||| || || || || ||||| |||||||||||||||| ||||| ||||
Db       7051 CATGGGTATTGTTAACTCCGAAAGCTAAGGATCAGTATGCAGTGGTTGGACACTCAGCAC 7110

Qy       1127 ACATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTC 1186
              |||||||||||||     |||||| |||||||| |||||||| ||||| ||||| |
Db       7111 ACATTGTTACACTGGCATCTGGCCGTGTGGTCATGTTGGTCATCTTCGGTCATTGCCCAC 7170

Qy       1187 TCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAA 1232
              |||||||||||||||  |||||||||||||| |||| || | |
Db       7171 TCTATGGATATATAAGCGTTGTGCAGGAATATGACTTGGGTATGTA 7216
```

RESULT 14
US-09-245-041-3/c
; Sequence 3, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 17056
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-3

Query Match 4.2%; Score 169.2; DB 3; Length 17056;
Best Local Similarity 86.9%; Pred. No. 8.3e-39;
Matches 186; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```
Qy        510 TTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAA 569
              ||  |||||||| |||||||| ||||| ||||||||||||||||||||||| ||||||
Db       4090 TTGTAGTTTTGACATGTGTCCGAATAATTGCTCAGGCCGAGGAGAGTGTAAGAGCAGTAA 4031

Qy        570 TAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACAT 629
              ||||||    ||||||| |||||||||||||||||||||||||||| || | |||||||||
Db       4030 CAGCAGCAGCGCTGTTGAGTGTGAATGTTCTGAAAACTGGAAAGGGGAGTCGTGTGACAT 3971

Qy        630 TCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGT 689
              |||||||||||||||||||||||| ||||||| |||||||||| ||| |||| |||
Db       3970 TCCTCACTGTACAGACAACTGTGGCTTTCCTCACCGAGGCATCTGTAATGCAAGCGATAC 3911

Qy        690 CAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGT 723
              |||||| |||||||||| | ||||||||||
Db       3910 CAGAGGGTGCTCCTGCTTTCCTCACTGGCAGGGT 3877
```

RESULT 15
US-09-245-041-6/c
; Sequence 6, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 226

; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-6

Query Match 3.2%; Score 132.2; DB 3; Length 226;
Best Local Similarity 78.6%; Pred. No. 3.6e-29;
Matches 158; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

```
Qy       2389 GATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCT 2448
             |||||||| ||||| |||||||||| || |||||||||| |||||||||||||||||||
Db        225 GATATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGTGATGCT 166

Qy       2449 GGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATC 2508
             || |||||||| || || ||||| || |||||||||||| | |||||||||||||||||||
Db        165 GGCTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGATTAAAGGCTGCAACCTGCATC 106

Qy       2509 AACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGG 2568
             ||||| |||||||| || ||||||||||||||||   ||  |||  | |    | |
Db        105 AACCCTCTCAATGGCAGCGTCTGTGAAAGGCCTGGTAAGGACATGGGTGCATATAGTGCT 46

Qy       2569 ACACCATGTGCCTTGAGGACA 2589
              ||  | | |||  ||   ||
Db         45 CCAGGAGGAGCCAAGACAGCA 25
```

Search completed: March 5, 2004, 12:20:15
Job time : 192.783 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 29, 2004, 14:49:01 ; Search time 11557 Seconds
(without alignments)
5359.280 Million cell updates/sec

Title: US-09-787-097-12
Perfect score: 7915
Sequence: 1 MVAAAAATEARLRRRTAATA.....KSGAVRNRKQQPPAQPGTCI 1429

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09787097/runat_26022004_062934_23420/app_query.fasta_1.1607
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09787097_@CGN_1_1_7794_@runat_26022004_062934_23420 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	7696.5	97.2	8589	6	AR164816	AR164816 Sequence
	2	7405	93.6	4493	4	AF531101	AF531101 Bos tauru
	3	7386	93.3	4284	10	AB062913	AB062913 Mesocrice
	4	7346.5	92.8	6083	10	AF116897	AF116897 Mus muscu
	5	7342.5	92.8	8739	10	AB038387	AB038387 Rattus no
	6	7303	92.3	8827	6	AR164807	AR164807 Sequence
	7	7302.5	92.3	4313	10	AF119821	AF119821 Mus muscu
	8	7089	89.6	3819	9	AF106861	AF106861 Homo sapi
	9	6870.5	86.8	4072	6	AR164817	AR164817 Sequence
	10	6612	83.5	3597	9	AF034957	AF034957 Homo sapi
	11	6509.5	82.2	4500	10	AB038388	AB038388 Rattus no
	12	5624	71.1	5683	10	AB093245	AB093245 Mus muscu
	13	4638.5	58.6	4140	9	AY442317	AY442317 Homo sapi
	14	4622	58.4	4463	9	AK127277	AK127277 Homo sapi
	15	4136.5	52.3	2625	6	AR164818	AR164818 Sequence
	16	3880	49.0	3609	6	AX503776	AX503776 Sequence
	17	3632.5	45.9	7302	9	AB011106	AB011106 Homo sapi
	18	3491.5	44.1	6370	6	AR164815	AR164815 Sequence
	19	3435	43.4	2419	6	AR164813	AR164813 Sequence
	20	3377.5	42.7	5002	10	BC027764	BC027764 Mus muscu
	21	3145	39.7	2578	9	AK000356	AK000356 Homo sapi
	22	3102.5	39.2	4997	10	BC050020	BC050020 Mus muscu
	23	2495	31.5	5632	9	AB011120	AB011120 Homo sapi
	24	2328	29.4	5271	3	AY051898	AY051898 Drosophil
c	25	2032.5	25.7	180499	3	AC007817	AC007817 Drosophil
c	26	2032.5	25.7	224910	3	AE003761	AE003761 Drosophil
c	27	1991	25.2	14345	2	AC013047	AC013047 Drosophil
	28	1939	24.5	3990	3	AF339882	AF339882 Caenorhab
	29	1825	23.1	1012	6	E26747	E26747 Structure a
	30	1797.5	22.7	3667	3	BT011125	BT011125 Drosophil
	31	1529.5	19.3	3297	10	BC030872	BC030872 Mus muscu
	32	1302.5	16.5	25305	3	CEF33C8	Z69790 Caenorhabdi
	33	1290	16.3	1818	9	BC047716	BC047716 Homo sapi
	34	1226.5	15.5	1051	6	AR164814	AR164814 Sequence
	35	1014	12.8	543	6	AX778545	AX778545 Sequence
	36	977.5	12.3	8875	6	BD185470	BD185470 Membrane
	37	977.5	12.3	9921	6	BD185469	BD185469 Membrane
	38	960.5	12.1	8589	6	AX921022	AX921022 Sequence
	39	950	12.0	6056	6	AX592784	AX592784 Sequence
	40	923	11.7	8307	9	AY280362	AY280362 Homo sapi
	41	922	11.6	6228	9	AB011541	AB011541 Homo sapi
	42	895	11.3	169427	3	AC007416	AC007416 Drosophil
	43	895	11.3	289516	3	AE003619	AE003619 Drosophil
	44	869.5	11.0	36993	2	AC018007	AC018007 Drosophil
	45	841.5	10.6	5693	10	BC036727	BC036727 Mus muscu

ALIGNMENTS

RESULT 1

```
AR164816
LOCUS       AR164816                8589 bp    DNA     linear   PAT 17-OCT-2001
DEFINITION  Sequence 14 from patent US 6274339.
ACCESSION   AR164816
VERSION     AR164816.1  GI:16238025
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 8589)
  AUTHORS   Moore,K. and Nagle,D.Lynn.
  TITLE     Methods and compositions for the diagnosis and treatment of body
            weight disorders, including obesity
  JOURNAL   Patent: US 6274339-A 14 14-AUG-2001;
FEATURES             Location/Qualifiers
     source          1. .8589
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:              0              Length:          8589
Score:                  7696.50        Matches:         1401
Percent Similarity:     97.91%         Conservative:    2
Best Local Similarity:  97.77%         Mismatches:      17
Query Match:            97.24%         Indels:          13
DB:                     6              Gaps:            3

US-09-787-097-12 (1-1429) x AR164816 (1-8589)

Qy          2 ValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAlaAla 21
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         10 GTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCG 69

Qy         22 LeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArgPro 41
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 CTCGCGGGCAGGAGCGGCGGGCCGCACCGACCCTGCACCGCGACAGGGGCCTGGAGGCCG 129

Qy         42 GlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSer----------------ProPro 56
              |||     |||   |||   |||||||||:::||||||                |||||
Db        130 GGACCGCGCGCCCGGCTGTGTCTCCCGCGGGTGCTGTCGCGGGCGCTGCCCCCGCCGCCG 189

Qy         57 LeuArgProArgLeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeu 76
              |||   |||   ||||||       ||||||||||||||||
Db        190 CTGCTGCCG---CTGCTCTTTTCGCTGCTGCTGCTGCCG---------------------- 225

Qy         77 LeuProCysGluAlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAla 96
              ||||||   ||||||||||||||||||   ||||||||||||||||||||||||||||||
Db        226 CTGCCCCGGGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCC 285

Qy         97 GluAlaLysGluCysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGly 116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 GAGGCCAAGGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGC 345

Qy        117 GlnCysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPhe 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 CAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTC 405

Qy        137 ArgLeuThrGlySerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThr 156
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        406 AGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACG 465

Qy        157 LysCysThrTrpLeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHis 176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        466 AAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCAT 525

Qy        177 PheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAla 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        526 TTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCA 585

Qy        197 ProLeuValAlaAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrVal 216
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        586 CCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTC 645

Qy        217 ProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyr 236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        646 CCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTAT 705

Qy        237 AsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGly 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        706 AATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGC 765

Qy        257 ArgGlyGluCysLysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsn 276
              ||||||||||||||||||||||||||||||:::||||||||||||||||||||||||||||
Db        766 CGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAAC 825

Qy        277 TrpLysGlyGluAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArg 296
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        826 TGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGA 885

Qy        297 GlyIleCysAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyPro 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        886 GGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCT 945

Qy        317 GlyCysSerValProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsn 336
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        946 GGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAAC 1005

Qy        337 LeuLysLeuProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValVal 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1006 TTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTT 1065

Qy        357 GlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSer 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1066 GGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCT 1125

Qy        377 ArgGluTrpLeuProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSer 396
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1126 AGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCT 1185

Qy        397 LeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsn 416
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1186 TTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAAT 1245

Qy        417 ValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrPro 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1246 GTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCT 1305

Qy        437 LysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsn 456
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1306 AAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAAT 1365

Qy        457 GlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsn 476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1366 GGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAAT 1425

Qy        477 ValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeu 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1426 GTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTT 1485

Qy        497 ValGlnGlyGlyTyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrVal 516
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1486 GTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTT 1545

Qy        517 HisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArg 536
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1546 CATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGA 1605

Qy        537 TyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeu 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1606 TATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTG 1665

Qy        557 HisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAsp 576
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1666 CACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGAC 1725
```

```
Qy       577 ThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAla 596
Db      1726 ACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCC 1785

Qy       597 CysAspArgTrpSerValLeuProArgProAspLeuHisHisAspValAsnArgPheGly 616
Db      1786 TGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGC 1845

Qy       617 HisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeu 636
Db      1846 CATTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTC 1905

Qy       637 LeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAla 656
Db      1906 CTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCT 1965

Qy       657 CysLeuAlaAlaGlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIle 676
Db      1966 TGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATC 2025

Qy       677 SerTrpAlaLeuAlaThrAspGluGlnGluGluLysLeuLysSerGluCysPheSerLys 696
Db      2026 TCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAA 2085

Qy       697 ArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsn 716
Db      2086 AGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAAC 2145

Qy       717 ThrAsnAspCysHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGlu 736
Db      2146 ACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAA 2205

Qy       737 GlyGlnIleSerIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCys 756
Db      2206 GGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGT 2265

Qy       757 AsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArg 776
Db      2266 AACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGG 2325

Qy       777 AsnGlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGly 796
Db      2326 AATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGA 2385

Qy       797 AsnSerCysLeuLysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCys 816
Db      2386 AACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGT 2445

Qy       817 ArgAsnHisAsnAlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeu 836
Db      2446 AGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTT 2505

Qy       837 LysGlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrp 856
Db      2506 AAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGG 2565

Qy       857 ValGlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThr 876
Db      2566 GTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACA 2625

Qy       877 AsnSerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeu 896
Db      2626 AATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTA 2685

Qy       897 SerGluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySer 916
Db      2686 TCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGT 2745

Qy       917 ValCysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArg 936
Db      2746 GTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGG 2805

Qy       937 ThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLys 956

Db      2806 ACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAG 2865

Qy       957 GlnCysValAspSerAsnAlaTyrValAlaSerProProGlyGlnCysMetGluTrp 976
Db      2866 CAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGG 2925

Qy       977 TyrThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCys 996
Db      2926 TATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGC 2985

Qy       997 LeuGluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCys 1016
Db      2986 TTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGC 3045

Qy      1017 IleGluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPhe 1036
Db      3046 ATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTC 3105

Qy      1037 TyrProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSer 1056
Db      3106 TATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCT 3165

Qy      1057 PheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIle 1076
Db      3166 TTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATC 3225

Qy      1077 CysGluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPhe 1096
Db      3226 TGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTC 3285

Qy      1097 TyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSer 1116
Db      3286 TACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCT 3345

Qy      1117 LeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGlu 1136
Db      3346 CTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAG 3405

Qy      1137 CysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyr 1156
Db      3406 TGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTAT 3465

Qy      1157 TyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyr 1176
Db      3466 TATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTAT 3525

Qy      1177 TyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPhe 1196
Db      3526 TACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTC 3585

Qy      1197 IleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGly 1216
Db      3586 ATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGA 3645

Qy      1217 ThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAsp 1236
Db      3646 ACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGAT 3705

Qy      1237 SerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyr 1256
Db      3706 AGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTAT 3765

Qy      1257 ValSerAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPhe 1276
Db      3766 GTCAGTAATTTCACCTGGCCCATCAAAATTCAGATTGCCTTCTCTCAGCACAGCAATTTT 3825

Qy      1277 MetAspLeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuLeuVal 1296
Db      3826 ATGGACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTG 3885

Qy      1297 AlaAlaValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeu 1316
```

```
Db       3886 GCTGCTGTGGTTTGGAAGATCAAACAAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTT 3945

Qy       1317 LeuArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGlu 1336
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3946 CTTCGAGAGATGCAACAGATGGCCAGCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAA 4005

Qy       1337 ThrAspGluGluProProAspLeuIleGlyGlySerIleLysThrValProLysProIle 1356
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4006 ACAGATGAGGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATT 4065

Qy       1357 AlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuPro 1376
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4066 GCACTGGAGCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCT 4125

Qy       1377 ArgGlyLeuGlyGlyIleProProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeu 1396
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4126 CGAGGCCTGGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTG 4185

Qy       1397 ValAspIleSerGlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAsn 1416
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4186 GTGGACATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAAC 4245

Qy       1417 ArgLysGlnGlnProProAlaGlnProGlyThrCysIle 1429
              |||||||||||||||||||||||||||||||||||||||
Db       4246 CGGAAGCAGCAGCCCCCTGCACAGCCTGGGACCTGCATC 4284
```

```
RESULT 2
AF531101
LOCUS       AF531101                4493 bp    mRNA    linear   MAM 19-AUG-2002
DEFINITION  Bos taurus attractin mRNA, complete cds.
ACCESSION   AF531101
VERSION     AF531101.1  GI:22297307
KEYWORDS    .
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 4493)
  AUTHORS   Graphodatskaya,D., Joerg,H. and Stranzinger,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2002) Animal Sciences, ETH, Tannenstrasse 1,
            Zurich, CH 8092, Switzerland
FEATURES             Location/Qualifiers
     source          1. .4493
                     /organism="Bos taurus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9913"
                     /chromosome="13"
     CDS             52. .4299
                     /note="transmembrane form; mahogany protein"
                     /codon_start=1
                     /product="attractin"
                     /protein_id="AAM95445.1"
                     /db_xref="GI:22297308"
                     /translation="MVAAATEARLRRTVVTPAPTSRNGRRCRDWDATRAGRPGLRTGL
                     RLPRRWRRLLLLLLPPLLLLPWAAEAAAAAAASVSGSAAAEAKECDRPCVNGGRCNPGT
                     GQCVCPAGWVGEQCQHCGGRFRLTGSSGFITDGPGNYKYKTKCTWLIEGQPNRIMRLR
                     FNHFATECSWDHLYVYDGDSIYAPLVAAFSGLIVPERDGNETVPEVVATSGYALLHFF
                     SDAAYNLTGFNITYNFDMCPNNCSGRGECKISNSSNTVQCECSENWKGEACDIPHCVN
                     NCGFPHRGICNSSDVRGCSCFSEWQGPGCSVPVPANQSFWTREEYSDLKLPRASHKAV
                     VNGNIMWVVGGYMFNHSDYNMVLAYDLASREWLALNRSVNSVVVRYGHSLALYQDKIY
                     MYGGKIDSTGNVTNELRVFHIHNESWVLLSPKAKEQYAVVGHSAHIVTLKNGRVVMLV
                     IFGHCPLYGYISSVQEYDLDKNTWSILHTQGALVQGGYGHSSVYDDRTKALYIHGGYK
                     AFSANKYRLADDLYRYDVDTQMWTILKDSRFFRYLHTAVIVSGTMLVFGGNTHNDTSM
                     SHGAKCFSSDFMAYDIDCDRWSVLPRPDLHHDVNRFGHSAVLYNRTMYVFGGFNSLLL
                     SDILVFTSEQCEAHQSEAACLAAGPGVRCMWDTGSSQCVSWELAPEAQEKIKSECFSK
                     RIFDHDRCDQLTDCYSCTANTNGCQWCNDHCVPTNHSCTEGQISIFKYDHCPKDNPMY
                     YCNKKTSCRSCALDQNCQWEPRNQECIALPENICGIGWHLVGNSCLKITTARETYDNA
                     KLSCRNHNAFLASLTTQKKVEFVLKQLRIMQSSQSMSKFTLTPWVGLRKINVSYWCWE
                     DMSPFTNSSLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLNGSVCERPANHSAKQ
                     CRTPCALRTACGECTSGSSECMWCSNMKQCVDSNAYVASFPFGQCMEWYTMSSCPPEN
                     CSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKMPSQGPTGNSYPQPLLNS
                     SMCLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCENLTTGKHCETCISGFYGDPTN
                     GGKCQPCRCNGHASLCNTNTGKCFCTTKGVKGDECQLCEVENRYQGNPLKGTCYYTLL
                     IDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASKNFNLNITWAASFSAGTQ
                     AGEEMPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTWPIKIQIAFSQHSNF
                     MDLVQFFVTPFSCFLSLLLVAAVVWKIKQSCWASRRREQLLREMQQMASRPFASVNVA
                     LETDEEPPDLIGGSIKTVPKPIALEPCFGNKAAVLSVFVRLPRGLGGIPPPGQSGLAV
                     ASALVDISQQMPVVYKEKSGAVRNRKQQPPAQPGTCI"
ORIGIN
```

```
Alignment Scores:
Pred. No.:              0               Length:         4493
Score:                  7405.00         Matches:        1340
Percent Similarity:     96.35%          Conservative:   33
Best Local Similarity:  94.04%          Mismatches:     40
Query Match:            93.56%          Indels:         12
DB:                     4               Gaps:           5

US-09-787-097-12 (1-1429) x AF531101 (1-4493)

Qy          5 AlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAlaAlaLeuAlaGly 24
              |||||||||||||||||||||||  |||||||||  |||  |||
Db         58 GCGGCGGCGACTGAGGCGAGGCTG---AGGAGGACCGTGGTGACACCGGCGCCCACGAGC 114

Qy         25 ArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArgProGlyLeuGly 44
              |||:::|||           |||||||  ||||||||||||||||||||||
Db        115 AGGAACGGCAGGCGGTGTCGGGACTGGGACGCGACTAGAGCTGGAAGGCCGGGGCTGCGG 174

Qy         45 AlaGlyLeuArgLeuProArgLeuLeuSerProProLeuArgProArgLeuLeuLeuLeu 64
              |||||||||||||||||  |||  |||  |||||||||
Db        175 ACCGGGCTGCGCCTCCCACGG----------------CGGTGGCGGCGCCTGCTGCTG 216

Qy         65 LeuLeuLeuLeuProProProLeuLeuLeuLeuLeuLeuProCysGluAlaGluAlaAla 84
              |||||||||      ||||||      ||||||||||||||      ||||||||||||
Db        217 CTGCTGCTC------CCGCCG------CTGCTGCTACTTCCCTGGGCGGCTGAGGCCGCG 264

Qy         85 AlaAlaAlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGluCysAspArgPro 104
              ||||||||||||:::||||||||||||||||||||||||||||||||||||||||||||
Db        265 GCGGCAGCGGCGTCGGTGTCGGGCTCAGCCGCAGCCGAGGCCAAGGAATGTGACCGGCCC 324

Qy        105 CysValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTrp 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 TGTGTCAACGGCGGCCGCTGCAACCCTGGCACCGGCCAGTGTGTCTGCCCCGCCGGCTGG 384

Qy        125 ValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPhe 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        385 GTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGATTT 444

Qy        145 ValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeuIleGluGly 164
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        445 ATAACAGATGGACCTGGAAATTATAAATATAAAACAAAATGCACATGGCTCATTGAAGGA 504

Qy        165 GlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAsp 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        505 CAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGAC 564

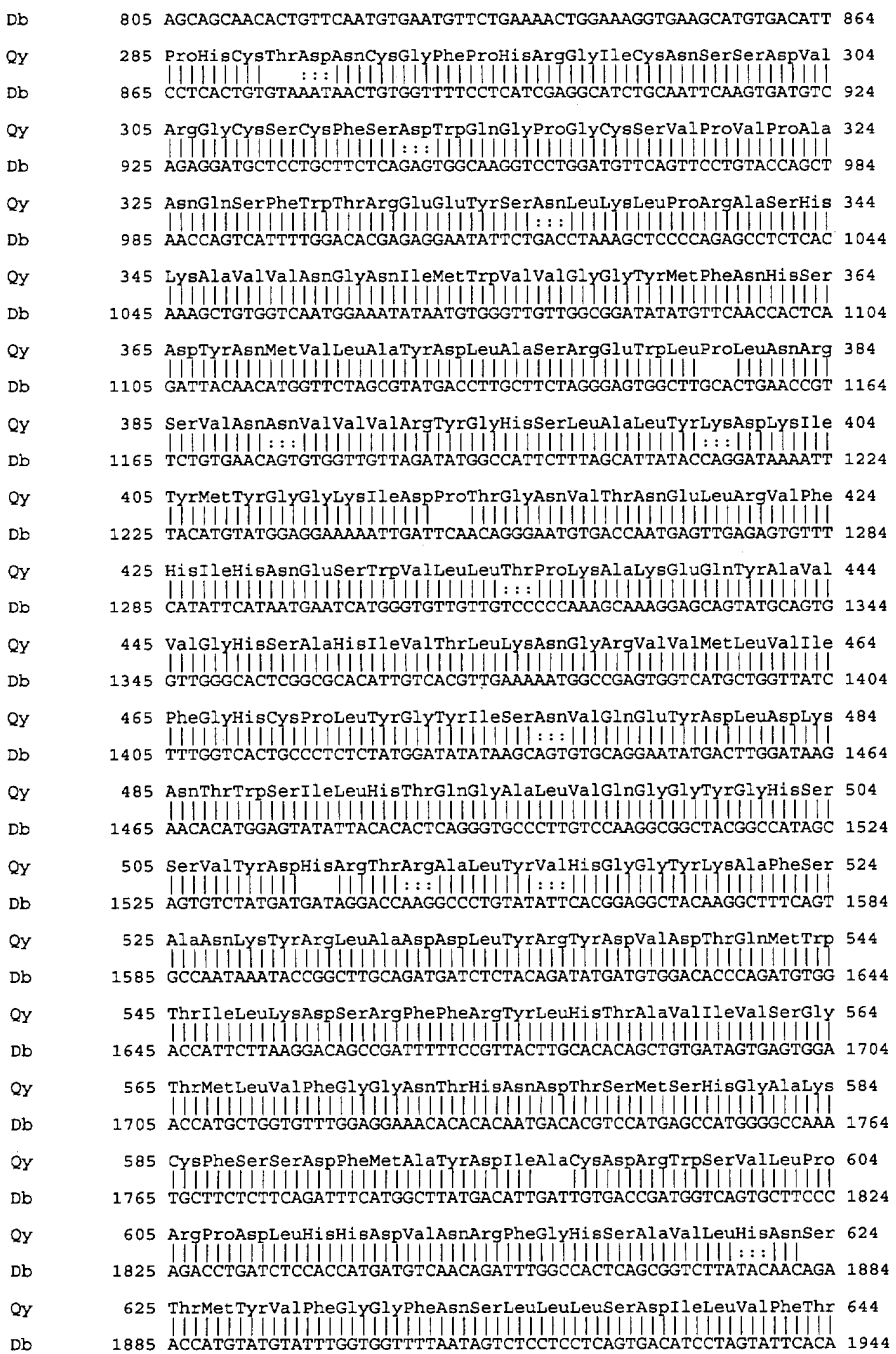
Qy        185 HisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaAlaPheSerGly 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        565 CATTTATATGTTTATGACGGGGACTCAATTTATGCACCCCTAGTTGCTGCCTTTAGCGGC 624

Qy        205 LeuIleValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGly 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        625 CTCATTGTTCCTGAGAGAGACGGCAATGAGACTGTCCCAGAGGTTGTCGCCACATCAGGT 684

Qy        225 TyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        685 TATGCCCTGCTGCATTTTTTTAGTGATGCTGCTTATAACTTGACTGGATTTAATATCACT 744

Qy        245 TyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsn 264
              |||:::||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        745 TACAATTTTGACATGTGTCCAAATAATTGCTCAGGCCGAGGAGAATGTAAGATCAGTAAT 804

Qy        265 SerSerGluThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIle 284
              ||||||  ||||||:::||||||||||||||||||||||||||||||||||||||||||
```


Db 805 AGCAGCAACACTGTTCAATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATT 864
Qy 285 ProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspVal 304
Db 865 CCTCACTGTGTAAATAACTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTC 924
Qy 305 ArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAla 324
Db 925 AGAGGATGCTCCTGCTTCTCAGAGTGGCAAGGTCCTGGATGTTCAGTTCCTGTACCAGCT 984
Qy 325 AsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHis 344
Db 985 AACCAGTCATTTTGGACACGAGAGGAATATTCTGACCTAAAGCTCCCCAGAGCCTCTCAC 1044
Qy 345 LysAlaValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHisSer 364
Db 1045 AAAGCTGTGGTCAATGGAAATATAATGTGGGTTGTTGGCGGATATATGTTCAACCACTCA 1104
Qy 365 AspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArg 384
Db 1105 GATTACAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTGCACTGAACCGT 1164
Qy 385 SerValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIle 404
Db 1165 TCTGTGAACAGTGTGGTTGTTAGATATGGCCATTCTTTAGCATTATACCAGGATAAAATT 1224
Qy 405 TyrMetTyrGlyGlyLysIleAspProThrGlyAsnValThrAsnGluLeuArgValPhe 424
Db 1225 TACATGTATGGAGGAAAAATTGATTCAACAGGGAATGTGACCAATGAGTTGAGAGTGTTT 1284
Qy 425 HisIleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaVal 444
Db 1285 CATATTCATAATGAATCATGGGTGTTGTTGTCCCCCAAAGCAAAGGAGCAGTATGCAGTG 1344
Qy 445 ValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValIle 464
Db 1345 GTTGGGCACTCGGCGCACATTGTCACGTTGAAAAATGGCCGAGTGGTCATGCTGGTTATC 1404
Qy 465 PheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLys 484
Db 1405 TTTGGTCACTGCCCTCTCTATGGATATATAAGCAGTGTGCAGGAATATGACTTGGATAAG 1464
Qy 485 AsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisSer 504
Db 1465 AACACATGGAGTATATTACACACTCAGGGTGCCCTTGTCCAAGGCGGCTACGGCCATAGC 1524
Qy 505 SerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheSer 524
Db 1525 AGTGTCTATGATGATAGGACCAAGGCCCTGTATATTCACGGAGGCTACAAGGCTTTCAGT 1584
Qy 525 AlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetTrp 544
Db 1585 GCCAATAAATACCGGCTTGCAGATGATCTCTACAGATATGATGTGGACACCCAGATGTGG 1644
Qy 545 ThrIleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerGly 564
Db 1645 ACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGA 1704
Qy 565 ThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLys 584
Db 1705 ACCATGCTGGTGTTTGGAGGAAACACACACAATGACACGTCCATGAGCCATGGGGCCAAA 1764
Qy 585 CysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuPro 604
Db 1765 TGCTTCTCTTCAGATTTCATGGCTTATGACATTGATTGTGACCGATGGTCAGTGCTTCCC 1824
Qy 605 ArgProAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSer 624
Db 1825 AGACCTGATCTCCACCATGATGTCAACAGATTTGGCCACTCAGCGGTCTTATACAACAGA 1884
Qy 625 ThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPheThr 644
Db 1885 ACCATGTATGTATTTGGTGGTTTTAATAGTCTCCTCCTCAGTGACATCCTAGTATTCACA 1944
Qy 645 SerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIle 664
Db 1945 TCGGAACAGTGTGAAGCACACCAGAGTGAAGCTGCTTGTCTCGCAGCAGGACCTGGTGTT 2004
Qy 665 ArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGlu 684
Db 2005 CGGTGTATGTGGGACACAGGGTCATCTCAGTGTGTCTCCTGGGAGCTGGCA---CCTGAA 2061
Qy 685 GlnGluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArgCys 704
Db 2062 GCACAAGAAAAGATAAAGTCAGAATGTTTTTCTAAAAGAATTTTTGACCATGACAGATGT 2121
Qy 705 AspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsn 724
Db 2122 GACCAGCTCACGGATTGTTACAGCTGCACAGCCAACACCAATGGCTGCCAGTGGTGCAAT 2181
Qy 725 AspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyr 744
Db 2182 GACCATTGTGTCCCTACGAACCACAGCTGCACAGAAGGCCAGATCTCCATTTTCAAGTAT 2241
Qy 745 GluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArgSer 764
Db 2242 GATCATTGCCCCAAGGATAACCCGATGTATTATTGTAACAAGAAGACCAGCTGCAGGAGC 2301
Qy 765 CysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuPro 784
Db 2302 TGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATCGCCCTGCCC 2361
Qy 785 GluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThrThr 804
Db 2362 GAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACT 2421
Qy 805 AlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAla 824
Db 2422 GCCAGGGAGACTTATGACAATGCAAAATTGTCCTGCAGAAACCACAATGCCTTTTTGGCT 2481
Qy 825 SerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGlnSer 844
Db 2482 TCTCTTACAACTCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATCATGCAGTCA 2541
Qy 845 SerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnVal 864
Db 2542 TCACAGAGCATGTCCAAGTTCACCTTAACCCCATGGGTTGGCCTTCGGAAGATCAATGTG 2601
Qy 865 SerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetPro 884
Db 2602 TCCTATTGGTGCTGGGAAGATATGTCCCCGTTCACAAATAGTTCACTACAGTGGATGCCA 2661
Qy 885 SerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeu 904
Db 2662 TCTGAGCCCAGTGATGCTGGATTCTGTGGGATCTTGTCAGAACCCAGTACTCGGGGCTTA 2721
Qy 905 LysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnHis 924
Db 2722 AAGGCTGCAACCTGCATCAATCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCTAACCAC 2781
Qy 925 SerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThrSer 944
Db 2782 AGTGCCAAGCAGTGTCGGACCCCTTGTGCCTTGCGGACAGCCTGTGGGGAGTGTACCAGT 2841
Qy 945 GlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTyr 964
Db 2842 GGCAGCTCGGAGTGCATGTGGTGCAGCAACATGAAGCAGTGCGTGGACTCCAATGCCTAC 2901
Qy 965 ValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSerThrCysProPro 984
Db 2902 GTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAGTGGTATACCATGAGCAGCTGCCCCCCT 2961
Qy 985 GluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrp 1004
Db 2962 GAAAATTGTTCAGGCTACTGTACCTGTAGTCACTGCTTGGAGCAACCAGGCTGTGGCTGG 3021

```
Qy       1005 CysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyPro 1024
Db       3022 TGCACTGATCCTAGCAATACGGGCAAAGGGAAATGCATAGAGGGCTCCTATAAAGGACCA 3081

Qy       1025 ValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSer 1044
Db       3082 GTGAAGATGCCTTCTCAGGGCCCTACAGGAAATTCCTACCCACAGCCTCTTCTCAATTCC 3141

Qy       1045 SerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGln 1064
Db       3142 AGCATGTGTCTGGAGGACAGCAGATATAACTGGTCTTTCATTCATTGTCCAGCTTGCCAG 3201

Qy       1065 CysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThr 1084
Db       3202 TGCAATGGCCACAGCAAATGCATTAATCAGAGTATCTGTGAGAAGTGTGAGAACCTGACC 3261

Qy       1085 ThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyGly 1104
Db       3262 ACGGGCAAGCACTGTGAGACCTGCATATCTGGCTTCTATGGTGATCCAACCAACGGGGGA 3321

Qy       1105 LysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLys 1124
Db       3322 AAATGTCAACCATGCAGGTGCAATGGACATGCGTCGCTATGCAATACCAACACGGGCAAG 3381

Qy       1125 CysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGluAsn 1144
Db       3382 TGCTTCTGCACCACCAAGGGCGTCAAGGGGGATGAATGTCAGCTATGTGAGGTAGAAAAT 3441

Qy       1145 ArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrGln 1164
Db       3442 CGATACCAAGGAAACCCTCTCAAAGGAACGTGTTATTATACTCTTCTTATTGACTATCAG 3501

Qy       1165 PheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIleAsnPheValAla 1184
Db       3502 TTCACCTTTAGCCTATCCCAGGAAGATGACCGCTATTACACAGCCATCAATTTTGTGGCG 3561

Qy       1185 ThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsn 1204
Db       3562 ACACCTGATGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAAAACTTCAAC 3621

Qy       1205 LeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMetPro 1224
Db       3622 CTCAACATCACCTGGGCTGCCAGCTTCTCAGCTGGAACCCAAGCTGGAGAAGAGATGCCT 3681

Qy       1225 ValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPheAsp 1244
Db       3682 GTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGAT 3741

Qy       1245 PheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpProIle 1264
Db       3742 TTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATC 3801

Qy       1265 LysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePheVal 1284
Db       3802 AAAATTCAGATTGCATTCTCCCAGCACAGCAATTTTATGGACCTAGTACAGTTCTTCGTG 3861

Qy       1285 ThrPhePheSerCysPheLeuSerLeuLeuLeuValAlaAlaValValTrpLysIleLys 1304
Db       3862 ACCTTCTTCAGTTGTTTCCTCTCTCTGCTCCTGGTGGCTGCTGTGGTTTGGAAGATCAAA 3921

Qy       1305 GlnSerCysTrpAlaSerArgArgArgGluGlnLeuLeuArgGluMetGlnGlnMetAla 1324
Db       3922 CAAAGTTGTTGGGCTTCCAGACGTAGAGAGCAACTTCTTCGAGAGATGCAGCAGATGGCC 3981

Qy       1325 SerArgProPheAlaSerValAsnValAlaLeuGluThrAspGluGluProProAspLeu 1344
Db       3982 AGTCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGAAGAGCCTCCTGATCTT 4041

Qy       1345 IleGlyGlySerIleLysThrValProLysProIleAlaLeuGluProCysPheGlyAsn 1364
Db       4042 ATTGGGGGGAGTATAAAGACTGTTCCCAAGCCCATTGCCCTGGAGCCGTGTTTTGGCAAC 4101

Qy       1365 LysAlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyGlyIleProPro 1384
Db       4102 AAAGCTGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCTCGGTGGAATCCCTCCT 4161

Qy       1385 ProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetPro 1404
Db       4162 CCTGGGCAGTCCGGGCTCGCAGTGGCCAGTGCCCTGGTGGACATTTCTCAGCAGATGCCA 4221

Qy       1405 IleValTyrLysGluLysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaGln 1424
Db       4222 GTCGTCTACAAAGAGAAGTCAGGAGCAGTGAGAAACCGAAAACAGCAGCCCCCTGCGCAG 4281

Qy       1425 ProGlyThrCysIle 1429
Db       4282 CCTGGGACCTGCATT 4296
```

RESULT 3
AB062913
LOCUS AB062913 4284 bp mRNA linear ROD 16-JAN-2002
DEFINITION Mesocricetus auratus Atrn mRNA for attractin, complete cds.
ACCESSION AB062913
VERSION AB062913.1 GI:16930100
KEYWORDS .
SOURCE Mesocricetus auratus (golden hamster)
ORGANISM Mesocricetus auratus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae; Mesocricetus.
REFERENCE 1
AUTHORS Kuramoto,T., Nomoto,T., Fujiwara,A., Mizutani,M., Sugimura,T. and Ushijima,T.
TITLE Insertional mutation of the Attractin gene in the black tremor hamster
JOURNAL Mamm. Genome 13 (1), 36-40 (2002)
MEDLINE 21635551
PUBMED 11773967
REFERENCE 2 (bases 1 to 4284)
AUTHORS Kuramoto,T. and Ushijima,T.
TITLE Direct Submission
JOURNAL Submitted (08-JUN-2001) Takashi Kuramoto, National Cancer Center Research Institute, Carcinogenesis Division; Tsukiji 5-1-1, Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramot@gan2.ncc.go.jp, URL:http://www.ncc.go.jp/research/rat-genome/, Tel:81-3-3542-2511, Fax:81-3-5565-1753)

```
FEATURES             Location/Qualifiers
     source          1. .4284
                     /organism="Mesocricetus auratus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10036"
     gene            1. .4284
                     /gene="Atrn"
     CDS             1. .4284
                     /gene="Atrn"
                     /codon_start=1
                     /product="attractin"
                     /protein_id="BAB72012.1"
                     /db_xref="GI:16930101"
                     /translation="MVAAAAATEALLRGSTAATAAPAGRKVEQYRLCAATGAWRPGPR
                     AGLCLPWVVSRALPPPLLLLLLFSLLLLPLPREAEAAAVSAAVSGSAAAEAKECDRP
                     CVNGGRCNPGTGQCVCPAGWVGEQCQHCGGRFRLTGSSGFVTDGPGNYKYKTKCTWLI
                     EGQPNRIMRLRFNHFATECSWDHLYVYDGDSIYAPLIAAFSGLIVPERDGNETAPEVT
                     VTSGYALLHFFSDAAYNLTGFNITYNFDMCPNNCSGRGECKSSNSSNTVECECSENWK
                     GESCDIPHCTDNCGFPHRGICNASDVRGCSCSPHWQGPGCSIPVPANQSFWIREEYSD
                     FKLPRASHKAVVNGNIMWVVGGYMFNHSDYSMVLAYDLASREWLPLNHSVNNVVVRYG
                     HSLALHKDKIYMYGGKIDSTGNVTNELRVFHIHNESWVLLTPKAKEQYAVVGHSAHIV
                     TLATGRVVMLVIFGHCPLYGYISIVQEYDLEKNTWSILQTQGALVQGGYGHSSVYDHR
                     TKALYVHGGYKAFSANKYRLADDLYRYDVDTQMWTILKDSRFFRYLHTAVIVSGTMLV
                     FGGNTHNDTSMSHGAKCFSSDFMAYDIACDRWSVLPRPDLHHDVNRFGHSAVLHNSTM
                     YVFGGFNSLLLSDVLVFTSEQCDAHHSEAACVAAGPGIRCLWDTVSSRCTSWELATEE
                     QAEKLKSECFSKRTLDHDRCDQHTDCYSCTANTNDCHWCNDHCVPVNHSCTEGQISIS
                     KYDNCPKDNPMYYCNKKTSCRSCALDQNCQWEPRNQECIALPENICGIGWHLVGNSCL
                     KITTAKENYDNAKLSCRNHNAFLASLTSQKKVEFVLKHLRLMQSSQSMSKLTLTPWVG
                     LRKINVSYWCWEDMSPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLNGS
                     VCERPANHSAKQCRTPCALRTACSECTSSSSECMWCSNMKQCVDSNAYVASFPFGQCM
```

```
                     EWYTMSSCPPENCSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPIRMPSQAS
                     TGNVYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCENLTTGKHCE
                     TCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGDECQLCEVENRYQG
                     NPLKGTCYYTLLIDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASKNFNLN
                     ITWATSFSAGTQAGEEMPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTWPI
                     KIQIAFSQHSNFMDLVQFFVTFFSCFLSLLLVAAVVWKIKQSCWASRRREQLLREMQQ
                     MASRPFASVNVALETDEEPPDLIGGSIKTVPKPIALEPCFGNKAAVLSVFVRLPRGLG
                     GIPPPGQSGLAVASALVDISQQMPIVYKEKSGAVRNRKQQPPAQPGTCI"
ORIGIN

Alignment Scores:
Pred. No.:                0                Length:          4284
Score:                    7386.00          Matches:         1338
Percent Similarity:       95.73%           Conservative:    30
Best Local Similarity:    93.63%           Mismatches:      59
Query Match:              93.32%           Indels:          2
DB:                       10               Gaps:            1

US-09-787-097-12 (1-1429) x AB062913 (1-4284)

Qy          1 MetValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTGGCGGCGGCGGCGGCGACTGAGGCGCTGCTGAGGGGGAGCACGGCGGCGACAGCT 60

Qy         21 AlaLeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArg 40
              |||   |||||||||             :::             |||   |||   |||
Db         61 GCGCCCGCGGGCAGGAAGGTCGAGCAGTACCGACTCTGCGCCGCGACTGGGGCCTGGAGG 120

Qy         41 ProGlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSerProProLeuArgProArg 60
              ||||||    |||||||||   ||||||  ::::::|||      |||   |||
Db        121 CCGGGGCCGCGCGCGGGATTGTGTCTCCCGTGGGTGGTGTCGCGGGCTCTGCCCCCGCCG 180

Qy         61 LeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeuProCysGlu 80
              ||||||||||||||||||||          ||||||||||||   ||||||  |||
Db        181 CTGCTGCTGCTGCTGCTGCTCTTT------TCGCTACTGCTGCTGCCGCTGCCCCGAGAG 234

Qy         81 AlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGlu 100
              ||||||||||||||   :::|||||||||||||||||||||||||||||||||||||||
Db        235 GCCGAGGCCGCGGCGGTGTCGGCGGCGGTGTCAGGCTCGGCCGCAGCCGAGGCCAAGGAA 294

Qy        101 CysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCys 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 TGTGACCGGCCGTGTGTCAACGGCGGCCGCTGCAACCCTGGCACTGGCCAGTGCGTCTGC 354

Qy        121 ProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGly 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 CCCGCGGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGA 414

Qy        141 SerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrp 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 TCTTCTGGATTCGTAACAGATGGACCTGGGAATTATAAGTATAAAACGAAGTGCACATGG 474

Qy        161 LeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 CTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAA 534

Qy        181 CysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAla 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        535 TGTAGCTGGGACCATTTATATGTTTATGACGGGGACTCAATTTATGCACCTCTGATTGCT 594

Qy        201 AlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGluValVal 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 GCCTTTAGTGGTCTCATTGTTCCTGAGAGAGATGGCAATGAGACAGCCCCAGAGGTCACT 654

Qy        221 AlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGly 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 GTCACTTCAGGTTATGCACTGCTGCATTTTTTCAGTGATGCTGCTTACAATCTGACTGGA 714

Qy        241 PheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCys 260
              ||||||||||||||:::||||||||||||||||||||||||||||||||||||||||||
Db        715 TTTAATATCACTTACAATTTTGACATGTGTCCGAATAATTGCTCAGGCCGAGGAGAATGT 774

Qy        261 LysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsnTrpLysGlyGlu 280
              |||   ||||||||||||   ||||||||||||||||||||||||||||||||||||||||
Db        775 AAGAGCAGTAACAGCAGCAATACTGTTGAGTGTGAATGTTCTGAAAACTGGAAAGGGGAA 834

Qy        281 AlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsn 300
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        835 TCGTGTGACATTCCCCATTGTACAGACAACTGTGGGTTTCCTCACCGAGGCATCTGTAAT 894

Qy        301 SerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerVal 320
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db        895 GCAAGTGACGTGAGAGGGTGTTCCTGCTCCCCGCACTGGCAGGGTCCTGGATGTTCAATT 954

Qy        321 ProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuPro 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||||
Db        955 CCTGTGCCAGCTAACCAGTCTTTCTGGATTCGAGAAGAATATTCCGATTTCAAGCTCCCC 1014

Qy        341 ArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMet 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1015 AGAGCCTCTCATAAAGCTGTGGTCAATGGGAATATAATGTGGGTTGTTGGTGGATATATG 1074

Qy        361 PheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeu 380
              ||||||||||||||||||:::||||||||||||||||||||||||||||||||||||||
Db       1075 TTCAACCACTCAGATTACAGCATGGTTTTAGCGTATGACCTGGCTTCTAGGGAATGGCTT 1134

Qy        381 ProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyr 400
              ||||||||  ||||||||||||||||||||||||||||||||||||||||||||:::|||
Db       1135 CCGCTAAACCATTCTGTGAACAATGTGGTTGTAAGATATGGTCATTCTTTGGCATTGCAT 1194

Qy        401 LysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsnValThrAsnGlu 420
              ||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db       1195 AAGGATAAAATCTACATGTATGGAGGAAAAATTGATTCAACAGGGAATGTGACCAATGAG 1254

Qy        421 LeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGlu 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1255 CTGAGAGTATTTCATATTCATAATGAATCATGGGTATTGTTAACACCAAAAGCAAAGGAG 1314

Qy        441 GlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValVal 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1315 CAGTATGCAGTAGTTGGACACTCAGCACATATTGTTACACTGGCAACCGGCCGTGTGGTC 1374

Qy        461 MetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyr 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1375 ATGTTGGTCATCTTTGGTCATTGCCCACTCTATGGATATATAAGCATTGTGCAGGAGTAT 1434

Qy        481 AspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGly 500
              ||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1435 GACTTAGAAAAGAACACATGGAGTATACTACAAACTCAGGGTGCTCTTGTGCAAGGAGGC 1494

Qy        501 TyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyr 520
              ||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||||
Db       1495 TATGGCCACAGCAGTGTTTATGATCACAGGACCAAGGCCCTGTATGTTCATGGTGGCTAC 1554

Qy        521 LysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAsp 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1555 AAGGCTTTTAGCGCCAATAAATACCGGCTTGCAGATGACCTCTACAGATATGATGTGGAT 1614

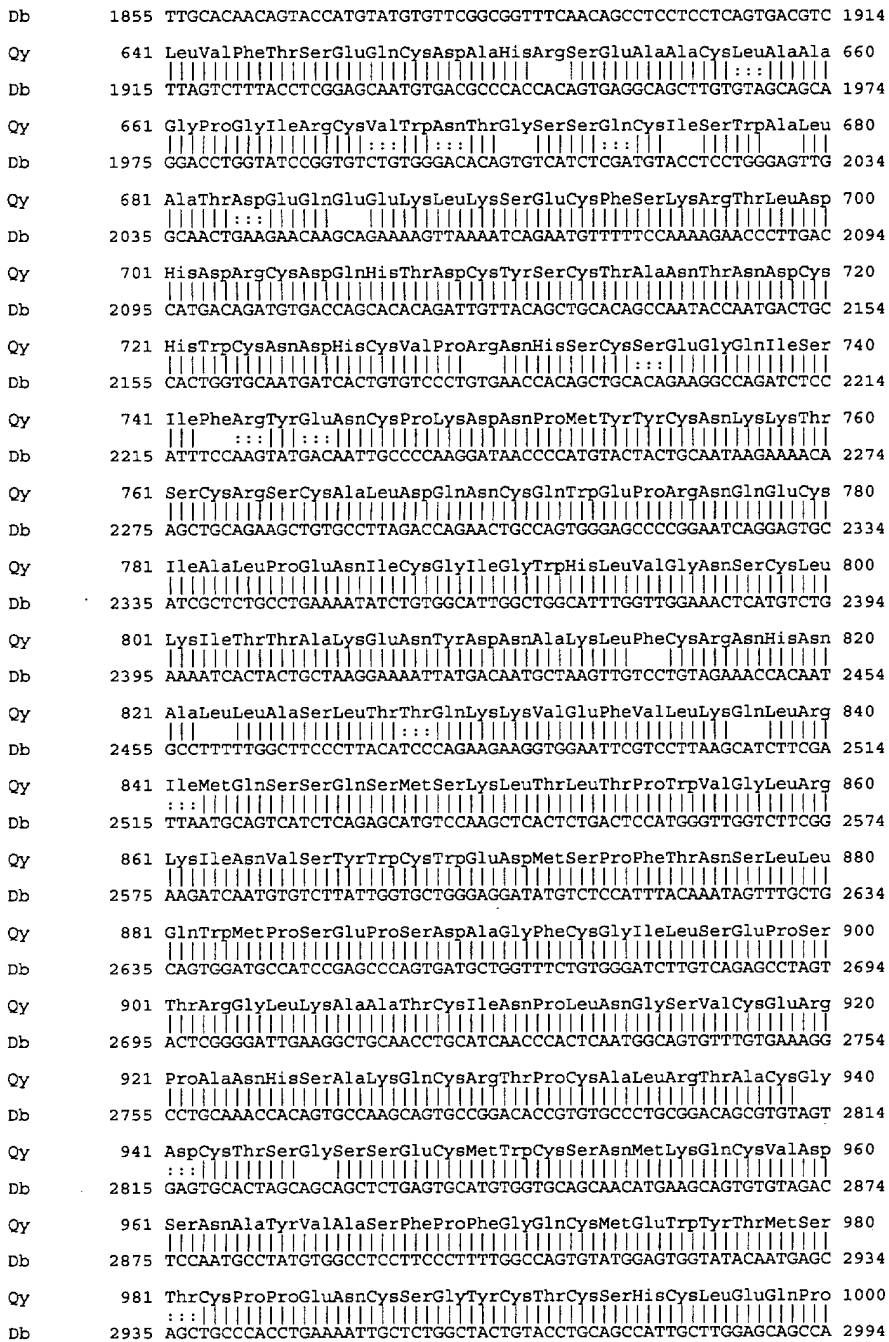
Qy        541 ThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaVal 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1615 ACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCATACAGCTGTG 1674

Qy        561 IleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSer 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1675 ATAGTGAGTGGAACCATGCTGGTGTTTGGAGGGAACACACACAACGACACATCCATGAGC 1734

Qy        581 HisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrp 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1735 CATGGCGCCAAATGCTTCTCCTCGGACTTCATGGCTTATGACATTGCTTGTGACCGATGG 1794

Qy        601 SerValLeuProArgProAspLeuHisHisAspValAsnArgPheGlyHisSerAlaVal 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1795 TCAGTGCTTCCCAGACCTGATCTCCACCACGATGTCAACAGATTTGGCCATTCAGCAGTC 1854

Qy        621 LeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIle 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
```


Db 1855 TTGCACAACAGTACCATGTATGTGTTCGGCGGTTTCAACAGCCTCCTCCTCAGTGACGTC 1914
Qy 641 LeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAla 660
Db 1915 TTAGTCTTTACCTCGGAGCAATGTGACGCCCACCACAGTGAGGCAGCTTGTGTAGCAGCA 1974
Qy 661 GlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeu 680
Db 1975 GGACCTGGTATCCGGTGTCTGTGGGACACAGTGTCATCTCGATGTACCTCCTGGGAGTTG 2034
Qy 681 AlaThrAspGluGlnGluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAsp 700
Db 2035 GCAACTGAAGAACAAGCAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACCCTTGAC 2094
Qy 701 HisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCys 720
Db 2095 CATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAATACCAATGACTGC 2154
Qy 721 HisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSer 740
Db 2155 CACTGGTGCAATGATCACTGTGTCCCTGTGAACCACAGCTGCACAGAAGGCCAGATCTCC 2214
Qy 741 IlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThr 760
Db 2215 ATTTCCAAGTATGACAATTGCCCCAAGGATAACCCCATGTACTACTGCAATAAGAAAACA 2274
Qy 761 SerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCys 780
Db 2275 AGCTGCAGAAGCTGTGCCTTAGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGC 2334
Qy 781 IleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeu 800
Db 2335 ATCGCTCTGCCTGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTCTG 2394
Qy 801 LysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsn 820
Db 2395 AAAATCACTACTGCTAAGGAAAATTATGACAATGCTAAGTTGTCCTGTAGAAACCACAAT 2454
Qy 821 AlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArg 840
Db 2455 GCCTTTTTGGCTTCCCTTACATCCCAGAAGAAGGTGGAATTCGTCCTTAAGCATCTTCGA 2514
Qy 841 IleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArg 860
Db 2515 TTAATGCAGTCATCTCAGAGCATGTCCAAGCTCACTCTGACTCCATGGGTTGGTCTTCGG 2574
Qy 861 LysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeu 880
Db 2575 AAGATCAATGTGTCTTATTGGTGCTGGGAGGATATGTCTCCATTTACAAATAGTTTGCTG 2634
Qy 881 GlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSer 900
Db 2635 CAGTGGATGCCATCCGAGCCCAGTGATGCTGGTTTCTGTGGGATCTTGTCAGAGCCTAGT 2694
Qy 901 ThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArg 920
Db 2695 ACTCGGGGATTGAAGGCTGCAACCTGCATCAACCCACTCAATGGCAGTGTTTGTGAAAGG 2754
Qy 921 ProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGly 940
Db 2755 CCTGCAAACCACAGTGCCAAGCAGTGCCGGACACCGTGTGCCCTGCGGACAGCGTGTAGT 2814
Qy 941 AspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAsp 960
Db 2815 GAGTGCACTAGCAGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTAGAC 2874
Qy 961 SerAsnAlaTyrValAlaSerProProGlyGlnCysMetGluTrpTyrThrMetSer 980
Db 2875 TCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAGTGGTATACAATGAGC 2934
Qy 981 ThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPro 1000
Db 2935 AGCTGCCCACCTGAAAATTGCTCTGGCTACTGTACCTGCAGCCATTGCTTGGAGCAGCCA 2994
Qy 1001 GlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySer 1020
Db 2995 GGCTGTGGCTGGTGTACTGATCCTAGCAATACTGGGAAAGGGAAATGTATCGAGGGCAGC 3054
Qy 1021 TyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnPro 1040
Db 3055 TATAAAGGACCCATAAGGATGCCTTCACAGGCTTCTACAGGAAATGTGTATCCACAGCCT 3114
Qy 1041 LeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCys 1060
Db 3115 CTTCTGAACTCTAGCATGTGTTTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGT 3174
Qy 1061 ProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCys 1080
Db 3175 CCAGCTTGCCAGTGCAATGGACACAGCAAATGTATTAACCAGAGTATCTGTGAGAAGTGT 3234
Qy 1081 GluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspPro 1100
Db 3235 GAGAACCTGACCACTGGCAAGCACTGTGAGACCTGCATATCTGGCTTCTATGGTGATCCG 3294
Qy 1101 ThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThr 1120
Db 3295 ACTAATGGAGGCAAATGTCAGCCATGTAAGTGCAACGGGCATGCATCGCTGTGCAATACC 3354
Qy 1121 AsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCys 1140
Db 3355 AACACCGGCAAGTGCTTCTGCACCACCAAGGGTGTCAAGGGAGATGAGTGCCAGCTATGT 3414
Qy 1141 GluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeu 1160
Db 3415 GAGGTAGAAAATCGATACCAAGGAAACCCTCTCAAAGGAACATGTTATTATACCCTTCTC 3474
Qy 1161 IleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIle 1180
Db 3475 ATTGACTATCAGTTCACTTTTAGCCTGTCCCAGGAAGACGATCGCTATTATACAGCCATC 3534
Qy 1181 AsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSer 1200
Db 3535 AATTTTGTGGCCACTCCTGATGAACAAAACAGGGATTTGGACATGTTTATCAATGCCTCA 3594
Qy 1201 LysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGly 1220
Db 3595 AAAAACTTCAACCTCAACATCACCTGGGCTACCAGCTTCTCAGCTGGAACCCAGGCAGGA 3654
Qy 1221 GluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsn 1240
Db 3655 GAAGAAATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGCTTCTCTAAT 3714
Qy 1241 GluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPhe 1260
Db 3715 GAGAAATTTGATTTTCGCAACCATCCAAACATTACTTTCTTTGTTTATGTCAGTAATTTC 3774
Qy 1261 ThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuVal 1280
Db 3775 ACCTGGCCCATCAAAATTCAGATTGCCTTCTCCCAGCACAGCAATTTTATGGACCTGGTA 3834
Qy 1281 GlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuLeuValAlaAlaValVal 1300
Db 3835 CAGTTCTTCGTGACTTTCTTCAGCTGTTTCCTCTCTCTGCTCCTAGTGGCTGCAGTGGTT 3894
Qy 1301 TrpLysIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeuLeuArgGluMet 1320
Db 3895 TGGAAGATCAAGCAGAGCTGCTGGGCATCCAGGAGGAGGGAGCAACTTCTTCGGGAGATG 3954
Qy 1321 GlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThrAspGluGlu 1340
Db 3955 CAGCAGATGGCCAGCCGTCCCTTTGCTTCTGTAAATGTTGCCTTGGAAACAGATGAAGAG 4014
Qy 1341 ProProAspLeuIleGlyGlySerIleLysThrValProLysProIleAlaLeuGluPro 1360
Db 4015 CCTCCTGATCTTATTGGGGGAAGTATAAAGACTGTCCCCAAGCCCATTGCCCTGGAGCCC 4074

```
Qy       1361 CysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGly 1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4075 TGCTTTGGTAACAAAGCTGCAGTCCTCTCTGTATTTGTGAGGCTCCCTCGAGGACTGGGT 4134

Qy       1381 GlyIleProProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSer 1400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4135 GGAATCCCTCCTCCTGGTCAGTCAGGTCTTGCTGTGGCCAGTGCCCTGGTGGACATTTCT 4194

Qy       1401 GlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAsnArgLysGlnGln 1420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4195 CAACAGATGCCGATAGTATACAAAGAGAAGTCAGGAGCGGTAAGAAACCGGAAGCAGCAG 4254

Qy       1421 ProProAlaGlnProGlyThrCysIle 1429
             |||||||||||||||||||||||||||||
Db       4255 CCCCCTGCACAGCCTGGAACCTGCATT 4281

RESULT 4
AF116897
LOCUS       AF116897               6083 bp    mRNA    linear   ROD 20-MAR-1999
DEFINITION  Mus musculus mahogany protein mRNA, complete cds.
ACCESSION   AF116897
VERSION     AF116897.1  GI:4454560
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 6083)
  AUTHORS   Nagle,D.L., McGrail,S.H., Vitale,J., Woolf,E.A., Dussault,B.J. Jr.,
            DiRocco,L., Holmgren,L., Montagno,J., Bork,P., Huszar,D.,
            Fairchild-Huntress,V., Ge,P., Keilty,J., Ebeling,C., Baldini,L.,
            Gilchrist,J., Burn,P., Carlson,G.A. and Moore,K.J.
  TITLE     The mahogany protein is a receptor involved in suppression of
            obesity
  JOURNAL   Nature 398 (6723), 148-152 (1999)
  MEDLINE   99184159
   PUBMED   10086355
REFERENCE   2  (bases 1 to 6083)
  AUTHORS   Nagle,D.L., McGrail,S.H., Vitale,J., Woolf,E.A., Dussault,B.J. Jr.,
            DiRocco,L., Holmgren,L., Montagno,J., Bork,P., Huszar,D.,
            Fairchild-Huntress,V., Ge,P., Keilty,J., Ebeling,C., Baldini,L.,
            Gilchrist,J., Burn,P., Carlson,G.A. and Moore,K.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-DEC-1998) Sequence Analysis, Millennium
            Pharmaceuticals, 640 Memorial Drive, Cambridge, MA 02139, USA
FEATURES             Location/Qualifiers
     source          1. .6083
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
     CDS             70. .4488
                     /note="MG"
                     /codon_start=1
                     /product="mahogany protein"
                     /protein_id="AAD20947.1"
                     /db_xref="GI:4454561"
                     /translation="MVAVAAAAATEARLRGSTTTTAAPAGRKGRQHRPCTATGAWRPG
                     PRARLCLPRVLSRALPPPPLLPLLFSLLLLPLPREAEAAAVAAAVSGSAAAEAKECDR
                     PCVNGGRCNPGTGQCVCPTGWVGEQCQHCGGRFRLTGSSGFVTDGPGNYKYKTKCTWL
                     IEGQPNRIMRLRFNHFATECSWDHLYVYDGDSIYAPLIAAFSGLIVPERDGNETAPEV
                     TVTSGYALLHFFSDAAYNLTGFNITYNFDMCPNNCSGRGECKSSNSSSAVECECSENW
                     KGESCDIPHCTDNCGFPHRGICNASDTRGCSCFPHWQGPGCSIPVPANQSFWTREEYS
                     DLKLPRASHKAVVNGNIMWVVGGYMFNHSDYSMVLAYDLTSREWLPLNHSVNSVVVRY
                     GHSLALHKDKIYMYGGKIDSTGNVTNELRVFHIHNESWVLLTPKAKDQYAVVGHSAHI
                     VTLASGRVVMLVIFGHCPLYGYISVVQEYDLEKNTWSILHTQGALVQGGYGHSSVYDD
                     RTKALYVHGGYKAFSANKYRLADDLYRYDVDTQMWTILKDSRFFRYLHTAVIVSGTML
                     VFGGNTHNDTSMSHGAKCFSSDFMAYDIACDRWSVLPRPELHHDVNRFGHSAVLYNST
                     MYVFGGFNSLLLSDVLVFTSEQCDAHRSEAACVAAGPGIRCLWDTQSSRCTSWELATE
                     EQAEKLKSECFSKRTLDHDRCDQHTDCYSCTANTNDCHWCNDHCVPVNHSCTEGQISI
                     AKYESCPKDNPMYYCNKKTSCRSCALDQNCQWEPRNQECIALPENICGNGWHLVGNSC
                     LKITTAKENYDNAKLSCRNHNAFLASLTSQKKVEFVLKQLRLMQSSQSMSKLTLTPWV
                     GLRKINVSYWCWEDMSPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLNG
                     SVCERPANHSAKQCRTPCALRTACGECTSSSSECMWCSNMKQCVDSNAYVASFPFGQC
                     MEWYTMSSCPPENCSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKMPSQA
                     SAGNVYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCEDLTTGKHC
                     ETCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGDECQLCEVENRYQ
                     GNPLKGTCYYTLLIDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASKNFNL
                     NITWATSFPAGTQTGEEVPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTWP
                     IKIQIAFSQHSNFMDLVQFFVTFFSCFLSLLLVAAVVWKIKQSCWASRRREQLLREMQ
                     QMASRPFASVNVALETDEEPPDLIGGSIKTVPKPIALEPCFGNKAAVLSVFVRLPRGL
                     GGIPPPGQSRSRCGQCPGGHFSADANSVQGEVRSCKKPEAAAAACTAWNLHLILGQEFA
                     LTEGVWSTKLSTGEAAGRKQAWGRRLGTLEASNSCA"
ORIGIN

Alignment Scores:
Pred. No.:                0                Length:          6083
Score:                    7346.50          Matches:         1332
Percent Similarity:       95.12%           Conservative:    32
Best Local Similarity:    92.89%           Mismatches:      56
Query Match:              92.82%           Indels:          14
DB:                       10               Gaps:            3

US-09-787-097-12 (1-1429) x AF116897 (1-6083)

Qy          2 ValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAlaAla 21
              ||||||||||||||||||||||||||||||||||||  |||     |||||||||
Db         79 GTGGCGGCGGCGGCGGCGACTGAGGCGCGGCTGAGGGGGAGCACGACGACGACAGCAGCG 138

Qy         22 LeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArgPro 41
              |||||||||  |||  |||  |||  |||  |||  |||  |||  |||||||||||||
Db        139 CCTGCGGGCAGGAAGGGCAGGCAGCACCGACCCTGCACCGCGACAGGGGCCTGGAGGCCG 198

Qy         42 GlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSer----------------ProPro 56
              |||   |||  |||  |||||||||:::||||||                  ||||||
Db        199 GGACCGCGCGCCCGGCTGTGTCTCCCGCGGGTGCTGTCGCGGGCGCTGCCCCCGCCGCCG 258

Qy         57 LeuArgProArgLeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeu 76
              |||  |||  ||||||      ||||||||||||||
Db        259 CTGCTGCCG---CTGCTCTTTTCGCTGCTGCTGCTGCCG---------------------- 294

Qy         77 LeuProCysGluAlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAla 96
              ||||||  ||||||||||||||||||  ||||||||||||||||||||||||||||||
Db        295 CTGCCCCGGGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCC 354

Qy         97 GluAlaLysGluCysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGly 116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 GAGGCCAAGGAATGTGACCGGCCGTGTGTCAACGGCGGCCGCTGCAACCCTGGCACCGGC 414

Qy        117 GlnCysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPhe 136
              ||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db        415 CAGTGCGTCTGCCCCACGGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTC 474

Qy        137 ArgLeuThrGlySerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThr 156
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 AGACTAACTGGCTCTTCTGGATTTGTAACAGATGGACCTGGGAATTATAAATATAAGACG 534

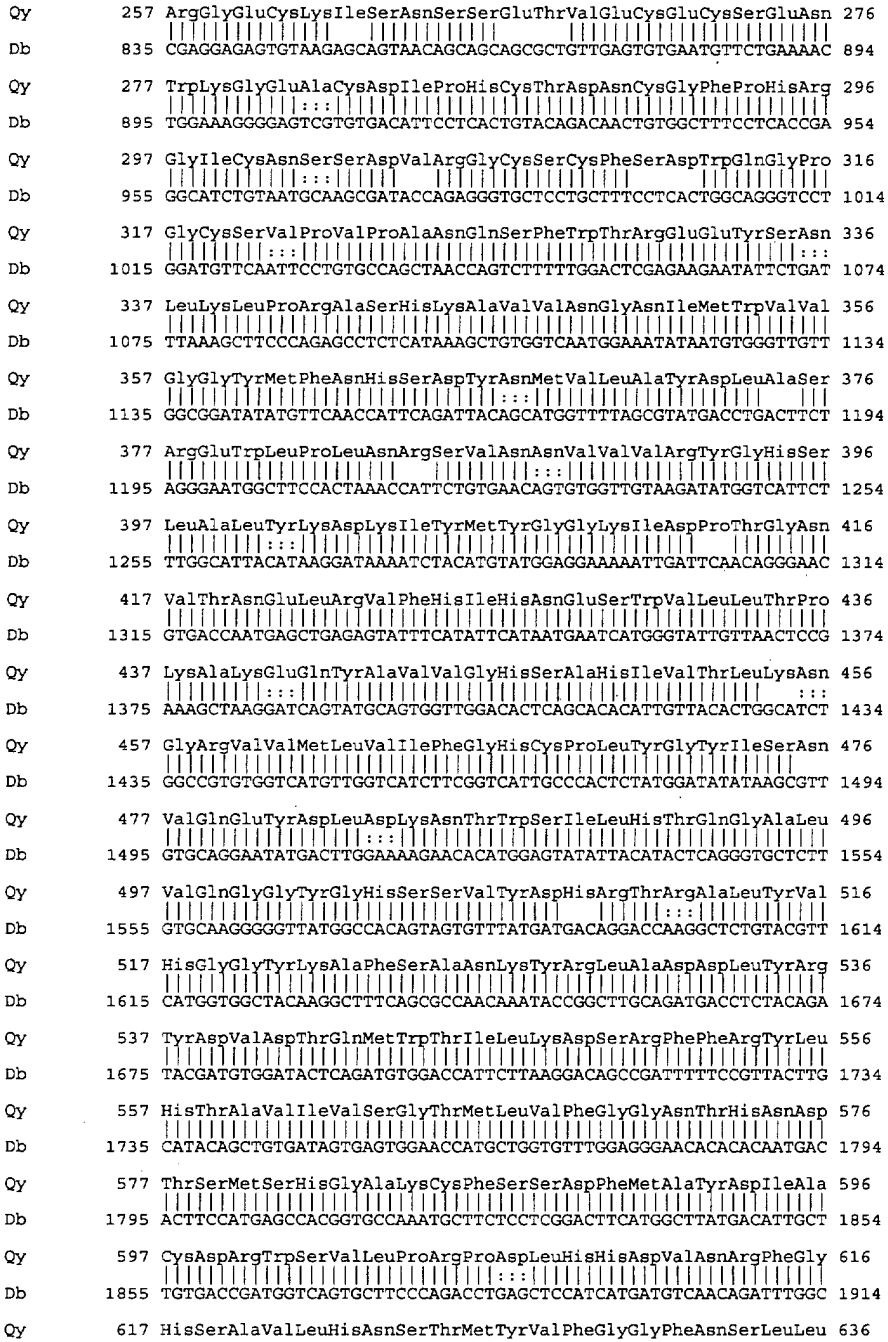
Qy        157 LysCysThrTrpLeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHis 176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        535 AAGTGCACATGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGCTTCAACCAT 594

Qy        177 PheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAla 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 TTTGCTACAGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATCTACGCA 654

Qy        197 ProLeuValAlaAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrVal 216
              ||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 CCTCTGATTGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCT 714

Qy        217 ProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyr 236
              |||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 CCTGAGGTCACTGTCACTTCAGGTTATGCACTGCTGCATTTTTTCAGTGATGCTGCTTAT 774

Qy        237 AsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGly 256
              ||||||||||||||||||||||||||:::||||||||||||||||||||||||||||||||
Db        775 AATCTGACTGGATTTAATATCACTTACAATTTTGACATGTGTCCGAATAATTGCTCAGGC 834
```


Qy 257 ArgGlyGluCysLysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsn 276
Db 835 CGAGGAGAGTGTAAGAGCAGTAACAGCAGCAGCGCTGTTGAGTGTGAATGTTCTGAAAAC 894
Qy 277 TrpLysGlyGluAlaCysAspIleProHisCysThrAspAsnCysGlyProHisArg 296
Db 895 TGGAAAGGGGAGTCGTGTGACATTCCTCACTGTACAGACAACTGTGGCTTTCCTCACCGA 954
Qy 297 GlyIleCysAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyPro 316
Db 955 GGCATCTGTAATGCAAGCGATACCAGAGGGTGCTCCTGCTTTCCTCACTGGCAGGGTCCT 1014
Qy 317 GlyCysSerValProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsn 336
Db 1015 GGATGTTCAATTCCTGTGCCAGCTAACCAGTCTTTTTGGACTCGAGAAGAATATTCTGAT 1074
Qy 337 LeuLysLeuProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValVal 356
Db 1075 TTAAAGCTTCCCAGAGCCTCTCATAAAGCTGTGGTCAATGGAAATATAATGTGGGTTGTT 1134
Qy 357 GlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSer 376
Db 1135 GGCGGATATATGTTCAACCATTCAGATTACAGCATGGTTTTAGCGTATGACCTGACTTCT 1194
Qy 377 ArgGluTrpLeuProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSer 396
Db 1195 AGGGAATGGCTTCCACTAAACCATTCTGTGAACAGTGTGGTTGTAAGATATGGTCATTCT 1254
Qy 397 LeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsn 416
Db 1255 TTGGCATTACATAAGGATAAAATCTACATGTATGGAGGAAAAATTGATTCAACAGGGAAC 1314
Qy 417 ValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrPro 436
Db 1315 GTGACCAATGAGCTGAGAGTATTTCATATTCATAATGAATCATGGGTATTGTTAACTCCG 1374
Qy 437 LysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsn 456
Db 1375 AAAGCTAAGGATCAGTATGCAGTGGTTGGACACTCAGCACACATTGTTACACTGGCATCT 1434
Qy 457 GlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsn 476
Db 1435 GGCCGTGTGGTCATGTTGGTCATCTTCGGTCATTGCCCACTCTATGGATATATAAGCGTT 1494
Qy 477 ValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeu 496
Db 1495 GTGCAGGAATATGACTTGGAAAAGAACACATGGAGTATATTACATACTCAGGGTGCTCTT 1554
Qy 497 ValGlnGlyGlyTyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrVal 516
Db 1555 GTGCAAGGGGGTTATGGCCACAGTAGTGTTTATGATGACAGGACCAAGGCTCTGTACGTT 1614
Qy 517 HisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArg 536
Db 1615 CATGGTGGCTACAAGGCTTTCAGCGCCAACAAATACCGGCTTGCAGATGACCTCTACAGA 1674
Qy 537 TyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeu 556
Db 1675 TACGATGTGGATACTCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTG 1734
Qy 557 HisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAsp 576
Db 1735 CATACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGAGGGAACACACACAATGAC 1794
Qy 577 ThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAla 596
Db 1795 ACTTCCATGAGCCACGGTGCCAAATGCTTCTCCTCGGACTTCATGGCTTATGACATTGCT 1854
Qy 597 CysAspArgTrpSerValLeuProArgProAspLeuHisHisAspValAsnArgPheGly 616
Db 1855 TGTGACCGATGGTCAGTGCTTCCCAGACCTGAGCTCCATCATGATGTCAACAGATTTGGC 1914
Qy 617 HisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeu 636
Db 1915 CATTCAGCAGTCTTGTACAACAGCACCATGTATGTGTTCGGCGGCTTCAACAGCCTCCTC 1974
Qy 637 LeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAla 656
Db 1975 CTCAGTGACGTCTTGGTCTTTACCTCGGAGCAGTGCGATGCACACCGCAGTGAAGCTGCT 2034
Qy 657 CysLeuAlaAlaGlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIle 676
Db 2035 TGTGTGGCAGCAGGACCTGGTATCCGGTGTCTGTGGGACACACAGTCGTCTCGATGTACC 2094
Qy 677 SerTrpAlaLeuAlaThrAspGluGlnGluGluLysLeuLysSerGluCysPheSerLys 696
Db 2095 TCCTGGGAGTTGGCAACTGAAGAACAAGCAGAAAAGTTAAAATCAGAGTGTTTTTCTAAA 2154
Qy 697 ArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsn 716
Db 2155 AGAACCCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAAT 2214
Qy 717 ThrAsnAspCysHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGlu 736
Db 2215 ACCAATGACTGCCACTGGTGCAATGATCACTGTGTCCCTGTGAACCACAGCTGCACAGAA 2274
Qy 737 GlyGlnIleSerIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCys 756
Db 2275 GGCCAGATCTCCATTGCCAAGTATGAGAGTTGCCCCAAGGATAACCCCATGTACTACTGC 2334
Qy 757 AsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArg 776
Db 2335 AATAAGAAAACCAGCTGCAGGAGCTGTGCCCTAGACCAGAACTGCCAGTGGGAGCCCCGG 2394
Qy 777 AsnGlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGly 796
Db 2395 AATCAAGAGTGCATCGCCCTGCCGGAAAATATCTGTGGCAATGGCTGGCATTTGGTTGGA 2454
Qy 797 AsnSerCysLeuLysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCys 816
Db 2455 AACTCGTGTCTGAAAATCACTACTGCTAAGGAGAATTATGACAATGCTAAATTGTCCTGT 2514
Qy 817 ArgAsnHisAsnAlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeu 836
Db 2515 AGGAACCACAATGCCTTTTTGGCTTCCCTCACATCCCAGAAGAAGGTGGAGTTTGTCCTT 2574
Qy 837 LysGlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrp 856
Db 2575 AAGCAGCTTCGATTAATGCAATCATCTCAAAGTATGTCCAAGCTCACTCTGACTCCATGG 2634
Qy 857 ValGlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThr 876
Db 2635 GTTGGTCTTCGGAAGATCAATGTGTCTTACTGGTGCTGGGAGGATATGTCTCCATTCACA 2694
Qy 877 AsnSerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeu 896
Db 2695 AATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGTGATGCTGGCTTCTGTGGGATCTTG 2754
Qy 897 SerGluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySer 916
Db 2755 TCAGAGCCTAGTACTCGGGGATTAAAGGCTGCAACCTGCATCAACCCTCTCAATGGCAGC 2814
Qy 917 ValCysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArg 936
Db 2815 GTCTGTGAAAGGCCTGCAAACCACAGTGCCAAGCAGTGCCGGACACCATGTGCCCTGCGG 2874
Qy 937 ThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLys 956
Db 2875 ACAGCGTGTGGCGAGTGCACTAGCAGCAGCTCGGAGTGCATGTGGTGCAGTAACATGAAG 2934
Qy 957 GlnCysValAspSerAsnAlaTyrValAlaSerProProGlyGlnCysMetGluTrp 976
Db 2935 CAGTGTGTGGACTCCAATGCCTACGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGG 2994
Qy 977 TyrThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCys 996

```
Db       2995 TATACGATGAGCAGCTGCCCACCTGAAAATTGCTCTGGCTACTGTACCTGCAGCCATTGC 3054
Qy        997 LeuGluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCys 1016
Db       3055 TTGGAGCAGCCAGGCTGTGGTTGGTGTACTGATCCTAGCAATACTGGGAAAGGAAAATGT 3114
Qy       1017 IleGluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPhe 1036
Db       3115 ATTGAGGGCAGCTATAAAGGACCTGTGAAGATGCCGTCACAGGCCTCTGCAGGAAATGTG 3174
Qy       1037 TyrProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSer 1056
Db       3175 TATCCACAGCCCCTTCTGAACTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCT 3234
Qy       1057 PheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIle 1076
Db       3235 TTCATTCACTGTCCAGCTTGCCAGTGCAACGGACACAGCAAATGCATCAACCAGAGTATC 3294
Qy       1077 CysGluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPhe 1096
Db       3295 TGTGAGAAGTGTGAGGACCTGACCACGGGCAAGCACTGCGAGACCTGCATATCTGGCTTC 3354
Qy       1097 TyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSer 1116
Db       3355 TATGGTGACCCGACTAATGGAGGCAAATGTCAGCCATGCAAGTGCAATGGGCACGCATCA 3414
Qy       1117 LeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGlu 1136
Db       3415 CTGTGCAACACCAACACCGGCAAGTGCTTCTGTACCACCAAAGGTGTCAAGGGGGACGAG 3474
Qy       1137 CysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyr 1156
Db       3475 TGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAAAGGAACATGCTAC 3534
Qy       1157 TyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyr 1176
Db       3535 TATACCCTTCTCATTGACTATCAGTTCACCTTTAGCCTGTCCCAGGAAGACGACCGCTAC 3594
Qy       1177 TyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPhe 1196
Db       3595 TACACAGCCATCAACTTTGTGGCTACTCCTGATGAACAAAACAGGGATTTGGACATGTTC 3654
Qy       1197 IleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGly 1216
Db       3655 ATCAATGCCTCCAAAAACTTCAACCTCAACATCACCTGGGCCACCAGCTTCCCAGCCGGA 3714
Qy       1217 ThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAsp 1236
Db       3715 ACCCAGACTGGAGAAGAGGTGCCTGTTGTTTCAAAAACCAACATCAAGGAATACAAAGAT 3774
Qy       1237 SerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyr 1256
Db       3775 AGCTTCTCTAATGAGAAATTTGATTTTCGCAACCATCCAAACATCACTTTCTTTGTTTAT 3834
Qy       1257 ValSerAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPhe 1276
Db       3835 GTCAGTAATTTCACTTGGCCCATCAAAATTCAGATTGCCTTCTCCCAGCACAGCAACTTC 3894
Qy       1277 MetAspLeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuLeuVal 1296
Db       3895 ATGGACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCGCTGCTTCTGGTG 3954
Qy       1297 AlaAlaValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeu 1316
Db       3955 GCTGCAGTGGTCTGGAAGATCAAGCAGAGCTGTTGGGCATCCAGGCGGAGAGAGCAACTT 4014
Qy       1317 LeuArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGlu 1336
Db       4015 CTTCGGGAGATGCAACAGATGGCCAGCCGCCCCTTTGCTTCTGTAAACGTTGCCTTGGAA 4074
Qy       1337 ThrAspGluGluProProAspLeuIleGlyGlySerIleLysThrValProLysProIle 1356
Db       4075 ACAGATGAAGAACCTCCTGATCTCATTGGGGGAAGTATAAAGACCGTTCCTAAGCCCATT 4134

Qy       1357 AlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuPro 1376
Db       4135 GCCCTGGAGCCCTGCTTTGGTAACAAAGCCGCAGTCCTCTCTGTATTCGTGAGGCTCCCT 4194
Qy       1377 ArgGlyLeuGlyGlyIleProProProGlyGlnSer-GlyLeuAlaValAlaSerAlaLe 1396
Db       4195 CGAGGACTGGGAGGAATCCCTCCTCCTGGTCAGTCAAGGTCTCGCTGTGGCCAGTGCCCT 4254
Qy       1396 uValAspIleSerGlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAs 1416
Db       4255 GGTGGACATTTCTCAGCAGATGCCAATAGTGTACAAGGAGAAGTCAGGAGCTGTAAGAAA 4314
Qy       1416 nArgLysGlnGlnProProAlaGlnProGlyThrCysIle 1429
Db       4315 CCGGAAGCAGCAGCCGCCTGCACAGCCTGGAACCTGCATT 4354
```

RESULT 5
AB038387

```
LOCUS       AB038387               8739 bp    mRNA    linear   ROD 18-JAN-2001
DEFINITION  Rattus norvegicus atrn mRNA for attractin, complete cds.
ACCESSION   AB038387
VERSION     AB038387.1  GI:12275307
KEYWORDS    attractin.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (sites)
  AUTHORS   Kuramoto,T., Kitada,K., Inui,T., Sasaki,Y., Ito,K., Hase,T.,
            Kawaguchi,S., Ogawa,Y., Nakao,K., Barsh,G.S., Nagao,M., Ushijima,T.
            and Serikawa,T.
  TITLE     Attractin/mahogany/zitter plays a critical role in myelination of
            the central nervous system
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (2), 559-564 (2001)
  MEDLINE   21143347
   PUBMED   11209055
REFERENCE   2  (bases 1 to 8739)
  AUTHORS   Kuramoto,T., Serikawa,T. and Ushijima,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-FEB-2000) Takashi Kuramoto, National Cancer Center
            Research Institute, Carcinogenesis Division; Tsukiji 5-1-1,
            Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramot@gan2.ncc.go.jp,
            Tel:81-33542-2511(ex.4521), Fax:81-35565-1753)
FEATURES             Location/Qualifiers
     source          1. .8739
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="WTC"
                     /db_xref="taxon:10116"
     gene            1. .8739
                     /gene="atrn"
     CDS             61. .4359
                     /gene="atrn"
                     /codon_start=1
                     /product="attractin"
                     /protein_id="BAB21017.1"
                     /db_xref="GI:12275308"
                     /translation="MVAAAAAAAEATEARLRGYTTATAAPAGWKERQHRPCAATGAWRP
                     WPRAGLCLPRVLSRALSPPPLLPLLPLLFSLLLLPLPREAEAAAVAAAVSGSAAAEAK
                     ECDRPCVNGGRCNPGTGQCVCPTGWVGEQCQHCGGRFRLTGSSGFVTDGPGNYKYKTK
                     CTWLIEGQPNKIMRLRFNHFATECSWDHLYVYDGDSIYAPLIAAFSGLIVPERDGNET
                     APEVTVTSGYALLHFFSDAAYNLTGFNITYNFDMCPNNCSGRGECKSSNSSSTVECEC
                     SENWKGESCDIPHCTDNCGFPHRGICNASDTRGCSCFPHWQGPGCSIPVPANQSFWTR
                     EEYSDLKLPRASHKAVVNGNIMWVVGGYMFNHSDYSMVLAYDLASREWLSLNHSVNSV
                     VVRYGHSLALHKDKIYMYGGKIDSTGNVTNELRVFHIHNESWVLLTPKAKDQYAVVGH
                     SAHIVTLSSGRVVMLVIFGHCPLYGYISVVQEYDLEKNTWSILQTQGALVQGGYGHSS
                     VYDHRTKALYVHGGYKAFSANKYRLADDLYRYHVDTQMWTILKDSRFFRYLHTAVIVS
                     GTMLVFGGNTHNDTSMSHGAKCFSSDFMAYDIACDRWSVLPRPELHHDVNRFGHSAVL
                     HNSTMYVFGGFNSLLLSDVLVFTSEQCDAHRSEAACVAAGPGIRCLWDTQSSRCTSWE
                     LATEEQAEKLKSECFSKRTLDHDRCDQHTDCYSCTANTNDCHWCNDHCVPVNHSCTEG
                     QISIAKYDNCPKDNPMYYCNKKTSCRSCALDQNCQWEPRNQECIALPENICGIGWHLV
```

GNSCLKITTAKENYDNAKLSCRNHNAFLASLTSQKKVEFVLKQLRLMQSSQSTSKLTL
TPWVGLRKINVSYWCWEDMSPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCIN
PLNGSVCERPANHSAKQCRTPCALRTACGECTSSSSECMWCSNMKQCVDSNAYVASFP
FGQCMEWYTMSSCPPENCSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKM
PSHASTGNVYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCEDLTT
GKHCETCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGEECQLCEVE
NRYQGNPLKGTCYYTLLIDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASK
NFNLNITWATSFPAGTQTGEEVPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSN
FTWPIKIQIAFSQHSNFMDLVQFFVTFFSCFLSLLLVAAVVWKIKQSCWASRRREQLL
REMQQMASRPFASVNVALETDEEPPDLIGGSIKTVPKPIALEPCFGNKAAVLSVFVRL
PRGLGGIPPPGQSGLAVASALVDISQQMPIVYKEKSGAVRNRKQQPPAQPGTCI"

ORIGIN

Alignment Scores:

Pred. No.:	0	Length:	8739
Score:	7342.50	Matches:	1331
Percent Similarity:	95.25%	Conservative:	33
Best Local Similarity:	92.95%	Mismatches:	65
Query Match:	92.77%	Indels:	3
DB:	10	Gaps:	1

US-09-787-097-12 (1-1429) x AB038387 (1-8739)

```
Qy          1 MetValAlaAlaAlaAla---------AlaThrGluAlaArgLeuArgArgArgThrAla 17
Db         61 ATGGTGGCGGCGGCGGCGGCGGCGGAGGCGACTGAGGCGCGGTTGAGGGGGTACACGACG 120

Qy         18 AlaThrAlaAlaLeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArg 37
Db        121 GCGACAGCAGCGCCCGCGGGCTGGAAGGAGAGGCAGCACCGACCCTGCGCCGCGACAGGG 180

Qy         38 AlaGlyArgProGlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSerProProLeu 57
Db        181 GCCTGGAGGCCGTGGCCGCGCGCCGGGCTGTGTCTCCCGCGGGTGCTGTCGCGGGCGCTA 240

Qy         58 ArgProArgLeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeuLeu 77
Db        241 TCCCCGCCGCCGCTGCTGCCGCTGCTGCCGCTGCTCTTTTCGCTATTGCTGCTGCCACTG 300

Qy         78 ProCysGluAlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAlaGlu 97
Db        301 CCCCGGGAGGCCGAGGCCGCTGCGGTGGCGGCGGCTGTGTCAGGCTCGGCCGCAGCTGAG 360

Qy         98 AlaLysGluCysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGlyGln 117
Db        361 GCCAAGGAATGTGACCGGCCGTGTGTCAACGGCGGCCGCTGCAACCCTGGCACCGGCCAG 420

Qy        118 CysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPheArg 137
Db        421 TGCGTCTGCCCCACGGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGA 480

Qy        138 LeuThrGlySerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLys 157
Db        481 CTAACTGGATCTTCTGGATTTGTAACAGATGGACCTGGGAATTATAAATATAAAACGAAA 540

Qy        158 CysThrTrpLeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPhe 177
Db        541 TGCACATGGCTCATTGAAGGACAGCCAAATAAGATAATGAGACTTCGATTCAACCATTTT 600

Qy        178 AlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaPro 197
Db        601 GCTACTGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATTTACGCACCT 660

Qy        198 LeuValAlaAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValPro 217
Db        661 CTGATTGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCCCCT 720

Qy        218 GluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsn 237
Db        721 GAAGTCACTGTCACTTCAGGTTATGCACTACTGCATTTTTTCAGTGATGCTGCTTATAAT 780

Qy        238 LeuThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArg 257
Db        781 CTGACTGGATTTAATATCACTTACAATTTTGACATGTGCCCGAATAATTGCTCAGGCCGA 840

Qy        258 GlyGluCysLysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsnTrp 277
Db        841 GGAGAATGTAAGAGCAGTAACAGCAGCAGCACTGTTGAGTGTGAATGTTCTGAAAACTGG 900

Qy        278 LysGlyGluAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGly 297
Db        901 AAAGGGGAGTCCTGTGACATTCCTCACTGTACAGACAACTGTGGCTTTCCTCACCGAGGC 960

Qy        298 IleCysAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyProGly 317
Db        961 ATCTGTAACGCAAGTGACACCAGAGGGTGCTCCTGCTTCCCTCACTGGCAGGGTCCTGGA 1020

Qy        318 CysSerValProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeu 337
Db       1021 TGTTCAATTCCTGTGCCAGCTAACCAGTCTTTTTGGACTCGAGAAGAATATTCTGACTTA 1080

Qy        338 LysLeuProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValValGly 357
Db       1081 AAGCTTCCCAGAGCCTCTCATAAAGCTGTGGTCAATGGAAATATAATGTGGGTTGTTGGC 1140

Qy        358 GlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArg 377
Db       1141 GGGTATATGTTCAACCATTCAGATTACAGCATGGTTTTAGCGTATGACCTGGCTTCTAGG 1200

Qy        378 GluTrpLeuProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSerLeu 397
Db       1201 GAATGGCTTTCACTAAACCATTCTGTGAACAGTGTGGTTGTAAGATATGGTCATTCTTTG 1260

Qy        398 AlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsnVal 417
Db       1261 GCATTACATAAGGATAAAATCTACATGTATGGAGGAAAAATTGATTCAACAGGGAATGTG 1320

Qy        418 ThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrProLys 437
Db       1321 ACCAACGAGCTGAGAGTATTTCATATTCATAATGAATCATGGGTATTGTTAACCCCCAAA 1380

Qy        438 AlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsnGly 457
Db       1381 GCTAAGGATCAGTATGCAGTGGTTGGACACTCAGCACACATTGTTACACTGTCATCTGGC 1440

Qy        458 ArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsnVal 477
Db       1441 CGTGTGGTCATGTTGGTCATCTTCGGTCATTGTCCGCTCTACGGATATATAAGTGTTGTG 1500

Qy        478 GlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuVal 497
Db       1501 CAGGAATATGACTTGGAAAAGAACACATGGAGTATATTACAAACTCAAGGTGCTCTAGTG 1560

Qy        498 GlnGlyGlyTyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrValHis 517
Db       1561 CAAGGGGGTTATGGCCACAGCAGTGTTTATGATCACAGGACCAAGGCTCTGTATGTTCAT 1620

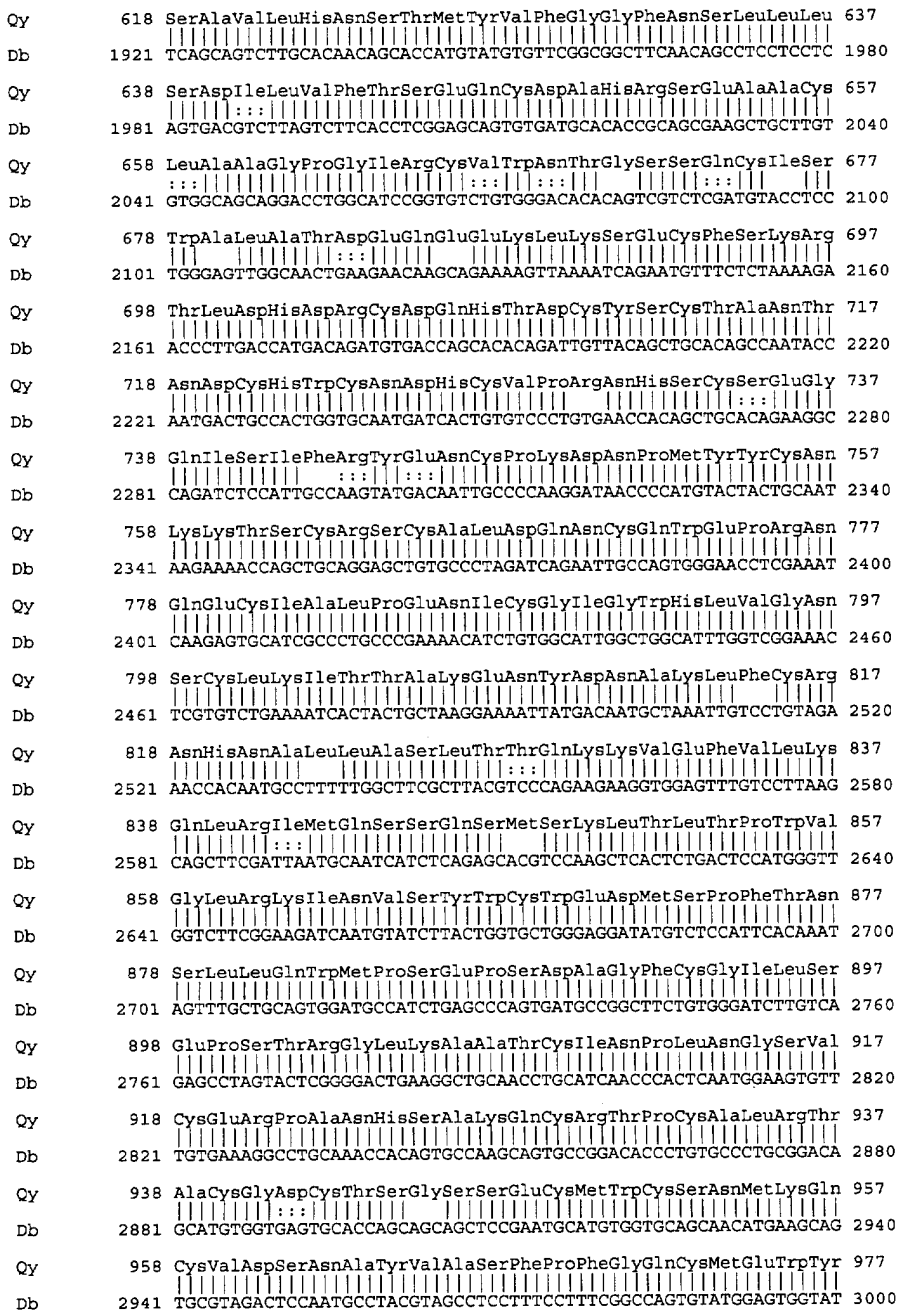
Qy        518 GlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyr 537
Db       1621 GGTGGCTACAAGGCTTTCAGCGCCAACAAGTACCGGCTTGCAGATGACCTCTACAGATAC 1680

Qy        538 AspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeuHis 557
Db       1681 CATGTGGACACTCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCAT 1740

Qy        558 ThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThr 577
Db       1741 ACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTCGGAGGGAACACACACAATGACACT 1800

Qy        578 SerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAlaCys 597
Db       1801 TCCATGAGCCACGGTGCCAAATGCTTCTCCTCAGACTTCATGGCTTATGACATTGCTTGT 1860

Qy        598 AspArgTrpSerValLeuProArgProAspLeuHisHisAspValAsnArgPheGlyHis 617
Db       1861 GACCGATGGTCAGTTCTCCCCAGACCTGAGCTCCATCACGATGTCAACCGATTTGGCCAC 1920
```


Qy 618 SerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeu 637
Db 1921 TCAGCAGTCTTGCACAACAGCACCATGTATGTGTTCGGCGGCTTCAACAGCCTCCTCCTC 1980
Qy 638 SerAspIleLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCys 657
Db 1981 AGTGACGTCTTAGTCTTCACCTCGGAGCAGTGTGATGCACACCGCAGCGAAGCTGCTTGT 2040
Qy 658 LeuAlaAlaGlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIleSer 677
Db 2041 GTGGCAGCAGGACCTGGCATCCGGTGTCTGTGGGACACACAGTCGTCTCGATGTACCTCC 2100
Qy 678 TrpAlaLeuAlaThrAspGluGlnGluGluLysLeuLysSerGluCysPheSerLysArg 697
Db 2101 TGGGAGTTGGCAACTGAAGAACAAGCAGAAAAGTTAAAATCAGAATGTTTCTCTAAAAGA 2160
Qy 698 ThrLeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThr 717
Db 2161 ACCCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAATACC 2220
Qy 718 AsnAspCysHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGluGly 737
Db 2221 AATGACTGCCACTGGTGCAATGATCACTGTGTCCCTGTGAACCACAGCTGCACAGAAGGC 2280
Qy 738 GlnIleSerIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsn 757
Db 2281 CAGATCTCCATTGCCAAGTATGACAATTGCCCCAAGGATAACCCCATGTACTACTGCAAT 2340
Qy 758 LysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsn 777
Db 2341 AAGAAAACCAGCTGCAGGAGCTGTGCCCTAGATCAGAATTGCCAGTGGGAACCTCGAAAT 2400
Qy 778 GlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsn 797
Db 2401 CAAGAGTGCATCGCCCTGCCCGAAAACATCTGTGGCATTGGCTGGCATTTGGTCGGAAAC 2460
Qy 798 SerCysLeuLysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArg 817
Db 2461 TCGTGTCTGAAAATCACTACTGCTAAGGAAAATTATGACAATGCTAAATTGTCCTGTAGA 2520
Qy 818 AsnHisAsnAlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLys 837
Db 2521 AACCACAATGCCTTTTTGGCTTCGCTTACGTCCCAGAAGAAGGTGGAGTTTGTCCTTAAG 2580
Qy 838 GlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpVal 857
Db 2581 CAGCTTCGATTAATGCAATCATCTCAGAGCACGTCCAAGCTCACTCTGACTCCATGGGTT 2640
Qy 858 GlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsn 877
Db 2641 GGTCTTCGGAAGATCAATGTATCTTACTGGTGCTGGGAGGATATGTCTCCATTCACAAAT 2700
Qy 878 SerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSer 897
Db 2701 AGTTTGCTGCAGTGGATGCCATCTGAGCCCAGTGATGCCGGCTTCTGTGGGATCTTGTCA 2760
Qy 898 GluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerVal 917
Db 2761 GAGCCTAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGAAGTGTT 2820
Qy 918 CysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThr 937
Db 2821 TGTGAAAGGCCTGCAAACCACAGTGCCAAGCAGTGCCGGACACCCTGTGCCCTGCGGACA 2880
Qy 938 AlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGln 957
Db 2881 GCATGTGGTGAGTGCACCAGCAGCAGCTCCGAATGCATGTGGTGCAGCAACATGAAGCAG 2940
Qy 958 CysValAspSerAsnAlaTyrValAlaSerProProGlyGlnCysMetGluTrpTyr 977
Db 2941 TGCGTAGACTCCAATGCCTACGTAGCCTCCTTTCCTTTCGGCCAGTGTATGGAGTGGTAT 3000
Qy 978 ThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeu 997
Db 3001 ACAATGAGCAGCTGCCCACCTGAAAATTGCTCTGGCTACTGTACCTGCAGCCATTGCTTG 3060
Qy 998 GluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIle 1017
Db 3061 GAGCAGCCAGGCTGTGGTTGGTGTACTGATCCTAGCAATACTGGGAAAGGCAAATGTATT 3120
Qy 1018 GluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyr 1037
Db 3121 GAGGGAAGCTATAAAGGACCCGTGAAGATGCCTTCACATGCCTCTACAGGAAATGTGTAC 3180
Qy 1038 ProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPhe 1057
Db 3181 CCACAGCCCCTTCTGAACTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTT 3240
Qy 1058 IleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCys 1077
Db 3241 ATTCACTGTCCAGCTTGCCAGTGCAATGGACACAGCAAGTGCATCAACCAGAGCATCTGT 3300
Qy 1078 GluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyr 1097
Db 3301 GAGAAGTGTGAGGACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTAT 3360
Qy 1098 GlyAspProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeu 1117
Db 3361 GGTGACCCGACTAATGGAGGCAAATGTCAGCCATGCAAGTGCAACGGGCACGCATCACTG 3420
Qy 1118 CysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCys 1137
Db 3421 TGCAACACCAACACTGGCAAGTGCTTCTGTACCACCAAGGGCGTCAAGGGAGAGGAGTGC 3480
Qy 1138 GlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyr 1157
Db 3481 CAGCTATGTGAGGTAGAAAATCGATACCAGGGAAACCCTCTCAAAGGAACATGTTACTAT 3540
Qy 1158 ThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyr 1177
Db 3541 ACCCTTCTCATTGACTATCAGTTCACTTTTAGCCTGTCCCAGGAAGACGACCGCTATTAC 3600
Qy 1178 ThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIle 1197
Db 3601 ACAGCCATCAATTTTGTGGCTACTCCTGATGAACAAAACAGGGATTTGGACATGTTCATC 3660
Qy 1198 AsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThr 1217
Db 3661 AATGCCTCCAAAAACTTCAACCTCAACATCACCTGGGCCACCAGTTTCCCAGCTGGAACC 3720
Qy 1218 GlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSer 1237
Db 3721 CAGACTGGAGAAGAGGTGCCTGTTGTTTCAAAAACCAACATCAAGGAATACAAAGATAGC 3780
Qy 1238 PheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrVal 1257
Db 3781 TTCTCTAATGAGAAATTTGATTTTCGCAACCATCCAAACATTACTTTCTTTGTTTATGTC 3840
Qy 1258 SerAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMet 1277
Db 3841 AGTAATTTCACTTGGCCCATCAAAATTCAAATTGCCTTCTCCCAGCACAGCAATTTTATG 3900
Qy 1278 AspLeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuLeuValAla 1297
Db 3901 GACCTGGTGCAGTTCTTCGTGACTTTCTTCAGCTGTTTCCTCTCTCTGCTTCTGGTGGCG 3960
Qy 1298 AlaValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeuLeu 1317
Db 3961 GCAGTGGTCTGGAAGATCAAGCAGAGCTGTTGGGCATCCAGGAGGAGAGAGCAACTTCTT 4020
Qy 1318 ArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThr 1337
Db 4021 CGGGAGATGCAACAGATGGCCAGCCGCCCTTTTGCTTCTGTAAATGTTGCCTTGGAAACA 4080
Qy 1338 AspGluGluProProAspLeuIleGlyGlySerIleLysThrValProLysProIleAla 1357

```
Db       4081 GATGAAGAACCTCCTGATCTCATTGGGGGAAGTATAAAGACTGTTCCTAAGCCCATTGCC 4140

Qy       1358 LeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProArg 1377

Db       4141 CTGGAGCCCTGCTTTGGTAACAAAGCTGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGA 4200

Qy       1378 GlyLeuGlyGlyIleProProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuVal 1397

Db       4201 GGACTGGGAGGAATCCCTCCTCCTGGTCAGTCAGGTCTTGCTGTGGCCAGTGCCCTGGTG 4260

Qy       1398 AspIleSerGlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAsnArg 1417

Db       4261 GACATTTCTCAGCAGATGCCAATAGTGTACAAGGAGAAGTCAGGAGCTGTAAGGAACCGG 4320

Qy       1418 LysGlnGlnProProAlaGlnProGlyThrCysIle 1429

Db       4321 AAGCAGCAGCCCCCTGCACAGCCTGGAACCTGCATT 4356

RESULT 6
AR164807
LOCUS       AR164807               8827 bp    DNA     linear   PAT 17-OCT-2001
DEFINITION  Sequence 1 from patent US 6274339.
ACCESSION   AR164807
VERSION     AR164807.1  GI:16238010
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 8827)
  AUTHORS   Moore,K. and Nagle,D.Lynn.
  TITLE     Methods and compositions for the diagnosis and treatment of body
            weight disorders, including obesity
  JOURNAL   Patent: US 6274339-A 1 14-AUG-2001;
FEATURES             Location/Qualifiers
     source          1. .8827
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:              0             Length:         8827
Score:                  7303.00       Matches:        1332
Percent Similarity:     90.99%        Conservative:   32
Best Local Similarity:  88.86%        Mismatches:     56
Query Match:            92.27%        Indels:         79
DB:                     6             Gaps:           4

US-09-787-097-12 (1-1429) x AR164807 (1-8827)

Qy          2 ValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAlaAla 21

Db         83 GTGGCGGCGGCGGCGGCGACTGAGGCGCGGCTGAGGGGGAGCACGAGGACGACAGCAGCG 142

Qy         22 LeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArgPro 41

Db        143 CCTGCGGGCAGGAAGGGCAGGCAGCACCGACCCTGCACCGCGACAGGGGCCTGGAGGCCG 202

Qy         42 GlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSer----------ProPro 56

Db        203 GGACCGCGCGCCCGGCTGTGTCTCCCGCGGGTGCTGTCGCGGGCGCTGCCCCCGCCGCCG 262

Qy         57 LeuArgProArgLeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeu 76

Db        263 CTGCTGCCG---CTGCTCTTTTCGCTGCTGCTGCTGCCG---------- 298

Qy         77 LeuProCysGluAlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAla 96

Db        299 CTGCCCCGGGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCC 358

Qy         97 GluAlaLysGluCysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGly 116

Db        359 GAGGCCAAGGAATGTGACCGGCCGTGTGTCAACGGCGGCCGCTGCAACCCTGGCACCGGC 418

Qy        117 GlnCysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPhe 136

Db        419 CAGTGCGTCTGCCCCACGGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTC 478

Qy        136 ---------- 136

Db        479 AGGACATCTGTCTCACGCCTATAATCACAGCTGTTCGGAAGGTGAGGCTGGAGGAACAGT 538

Qy        136 ---------- 136

Db        539 TCGAGGCAAGCTTCGGCTACAGAATAAGTTCAAGAGTAACCTGGGGCAACTTGGGCTTGT 598

Qy        136 ---------- 136

Db        599 CTCCAAAACCAAAATGAGCGAAAAGGAGCAAGCTAGAGTCTTTTGGGAAAATTTTAGCTG 658

Qy        137 ----------Arg-LeuThrGlySerSerGlyPheValThrAspGlyProGlyAs 151

Db        659 ACTAATTTTTCACCGAGAACTAACTGGCTCTTCTGGATTTGTAACAGATGGACCTGGGAA 718

Qy        151 nTyrLysTyrLysThrLysCysThrTrpLeuIleGluGlyGlnProAsnArgIleMetAr 171

Db        719 TTATAAATATAAGACGAAGTGCACATGGCTCATTGAAGGACAGCCAAATAGAATAATGAG 778

Qy        171 gLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGl 191

Db        779 ACTTCGCTTCAACCATTTTGCTACAGAATGTAGCTGGGACCATTTATATGTTTATGATGG 838

Qy        191 yAspSerIleTyrAlaProLeuValAlaAlaPheSerGlyLeuIleValProGluArgAs 211

Db        839 GGACTCAATCTACGCACCTCTGATTGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGA 898

Qy        211 pGlyAsnGluThrValProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePh 231

Db        899 TGGCAATGAGACGGCTCCTGAGGTCACTGTCACTTCAGGTTATGCACTGCTGCATTTTTT 958

Qy        231 eSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysPr 251

Db        959 CAGTGATGCTGCTTATAATCTGACTGGATTTAATATCACTTACAATTTTGACATGTGTCC 1018

Qy        251 oAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSerSerGluThrValGluCy 271

Db       1019 GAATAATTGCTCAGGCCGAGGAGAGTGTAAGAGCAGTAACAGCAGCAGCGCTGTTGAGTG 1078

Qy        271 sGluCysSerGluAsnTrpLysGlyGluAlaCysAspIleProHisCysThrAspAsnCy 291

Db       1079 TGAATGTTCTGAAAACTGGAAAGGGGAGTCGTGTGACATTCCTCACTGTACAGACAACTG 1138

Qy        291 sGlyPheProHisArgGlyIleCysAsnSerSerAspValArgGlyCysSerCysPheSe 311

Db       1139 TGGCTTTCCTCACCGAGGCATCTGTAATGCAAGCGATACCAGAGGGTGCTCCTGCTTTCC 1198

Qy        311 rAspTrpGlnGlyProGlyCysSerValProValProAlaAsnGlnSerPheTrpThrAr 331

Db       1199 TCACTGGCAGGGTCCTGGATGTTCAATTCCTGTGCCAGCTAACCAGTCTTTTTGGACTCG 1258

Qy        331 gGluGluTyrSerAsnLeuLysLeuProArgAlaSerHisLysAlaValValAsnGlyAs 351

Db       1259 AGAAGAATATTCTGATTTAAAGCTTCCCAGAGCCTCTCATAAAGCTGTGGTCAATGGAAA 1318

Qy        351 nIleMetTrpValValGlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAl 371

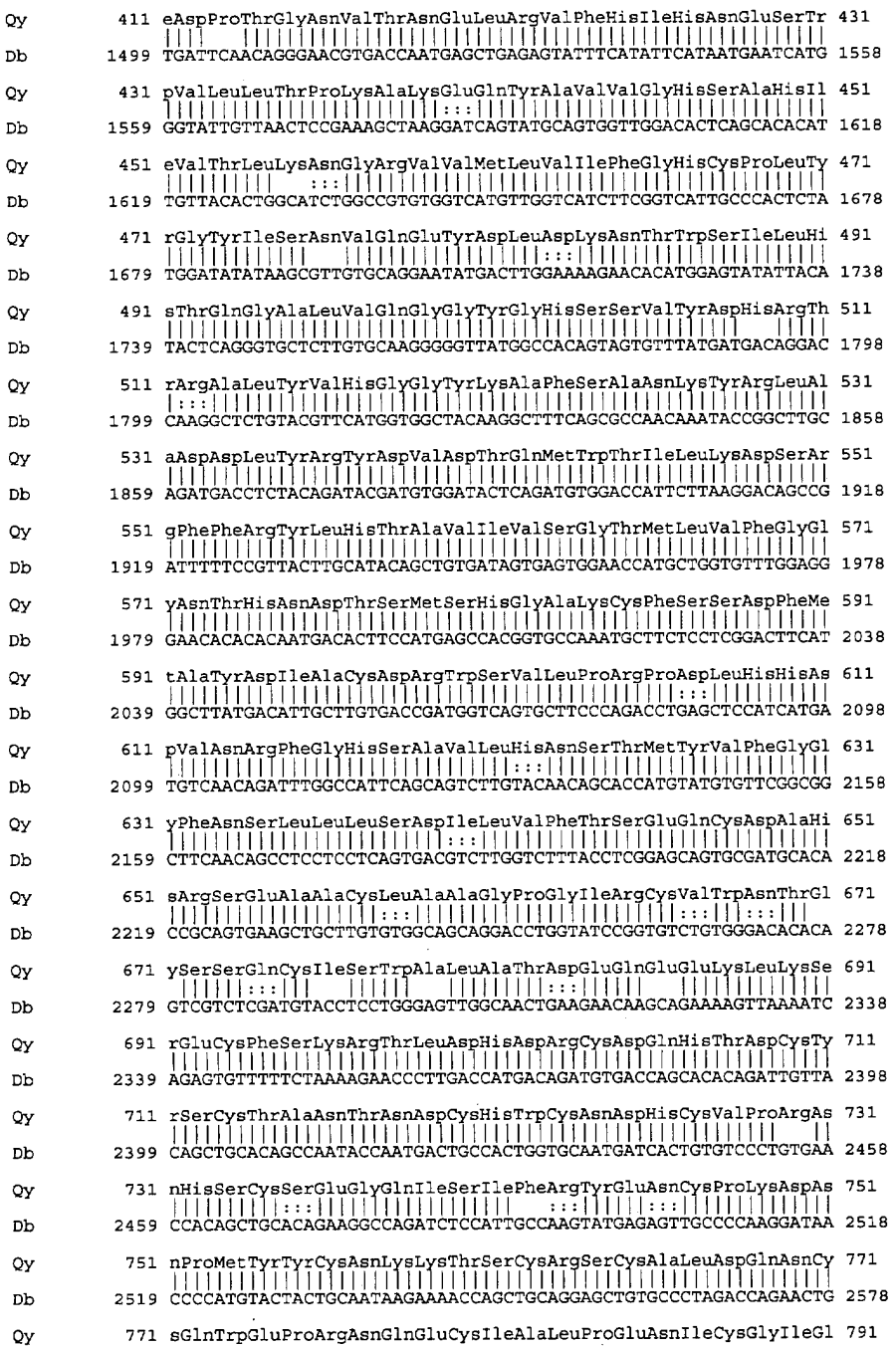
Db       1319 TATAATGTGGGTTGTTGGCGGATATATGTTCAACCATTCAGATTACAGCATGGTTTTAGC 1378

Qy        371 aTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSerValAsnAsnValValVa 391

Db       1379 GTATGACCTGACTTCTAGGGAATGGCTTCCACTAAACCATTCTGTGAACAGTGTGGTTGT 1438

Qy        391 lArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLysIl 411

Db       1439 AAGATATGGTCATTCTTTGGCATTACATAAGGATAAAATCTACATGTATGGAGGAAAAAT 1498
```


Qy 411 eAspProThrGlyAsnValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTr 431
Db 1499 TGATTCAACAGGGAACGTGACCAATGAGCTGAGAGTATTTCATATTCATAATGAATCATG 1558
Qy 431 pValLeuLeuThrProLysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIl 451
Db 1559 GGTATTGTTAACTCCGAAAGCTAAGGATCAGTATGCAGTGGTTGGACACTCAGCACACAT 1618
Qy 451 eValThrLeuLysAsnGlyArgValValMetLeuValIlePheGlyHisCysProLeuTy 471
Db 1619 TGTTACACTGGCATCTGGCCGTGTGGTCATGTTGGTCATCTTCGGTCATTGCCCACTCTA 1678
Qy 471 rGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHi 491
Db 1679 TGGATATATAAGCGTTGTGCAGGAATATGACTTGGAAAAGAACACATGGAGTATATTACA 1738
Qy 491 sThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisSerSerValTyrAspHisArgTh 511
Db 1739 TACTCAGGGTGCTCTTGTGCAAGGGGGTTATGGCCACAGTAGTGTTTATGATGACAGGAC 1798
Qy 511 rArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAl 531
Db 1799 CAAGGCTCTGTACGTTCATGGTGGCTACAAGGCTTTCAGCGCCAACAAATACCGGCTTGC 1858
Qy 531 aAspAspLeuTyrArgTyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerAr 551
Db 1859 AGATGACCTCTACAGATACGATGTGGATACTCAGATGTGGACCATTCTTAAGGACAGCCG 1918
Qy 551 gPhePheArgTyrLeuHisThrAlaValIleValSerGlyThrMetLeuValPheGlyGl 571
Db 1919 ATTTTTCCGTTACTTGCATACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGAGG 1978
Qy 571 yAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMe 591
Db 1979 GAACACACACAATGACACTTCCATGAGCCACGGTGCCAAATGCTTCTCCTCGGACTTCAT 2038
Qy 591 tAlaTyrAspIleAlaCysAspArgTrpSerValLeuProArgProAspLeuHisHisAs 611
Db 2039 GGCTTATGACATTGCTTGTGACCGATGGTCAGTGCTTCCCAGACCTGAGCTCCATCATGA 2098
Qy 611 pValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGl 631
Db 2099 TGTCAACAGATTTGGCCATTCAGCAGTCTTGTACAACAGCACCATGTATGTGTTCGGCGG 2158
Qy 631 yPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHi 651
Db 2159 CTTCAACAGCCTCCTCCTCAGTGACGTCTTGGTCTTTACCTCGGAGCAGTGCGATGCACA 2218
Qy 651 sArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgCysValTrpAsnThrGl 671
Db 2219 CCGCAGTGAAGCTGCTTGTGTGGCAGCAGGACCTGGTATCCGGTGTCTGTGGGACACACA 2278
Qy 671 ySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnGluGluLysLeuLysSe 691
Db 2279 GTCGTCTCGATGTACCTCCTGGGAGTTGGCAACTGAAGAACAAGCAGAAAAGTTAAAATC 2338
Qy 691 rGluCysPheSerLysArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTy 711
Db 2339 AGAGTGTTTTTCTAAAAGAACCCTTGACCATGACAGATGTGACCAGCACACAGATTGTTA 2398
Qy 711 rSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAspHisCysValProArgAs 731
Db 2399 CAGCTGCACAGCCAATACCAATGACTGCCACTGGTGCAATGATCACTGTGTCCCTGTGAA 2458
Qy 731 nHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGluAsnCysProLysAspAs 751
Db 2459 CCACAGCTGCACAGAAGGCCAGATCTCCATTGCCAAGTATGAGAGTTGCCCCAAGGATAA 2518
Qy 751 nProMetTyrTyrCysAsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCy 771
Db 2519 CCCCATGTACTACTGCAATAAGAAAACCAGCTGCAGGAGCTGTGCCCTAGACCAGAACTG 2578
Qy 771 sGlnTrpGluProArgAsnGlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGl 791
Db 2579 CCAGTGGGAGCCCCGGAATCAAGAGTGCATCGCCCTGCCGGAAAATATCTGTGGCAATGG 2638
Qy 791 yTrpHisLeuValGlyAsnSerCysLeuLysIleThrThrAlaLysGluAsnTyrAspAs 811
Db 2639 CTGGCATTTGGTTGGAAACTCGTGTCTGAAAATCACTACTGCTAAGGAGAATTATGACAA 2698
Qy 811 nAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSerLeuThrThrGlnLysLy 831
Db 2699 TGCTAAATTGTCCTGTAGGAACCACAATGCCTTTTTGGCTTCCCTCACATCCCAGAAGAA 2758
Qy 831 sValGluPheValLeuLysGlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLe 851
Db 2759 GGTGGAGTTTGTCCTTAAGCAGCTTCGATTAATGCAATCATCTCAAAGTATGTCCAAGCT 2818
Qy 851 uThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAs 871
Db 2819 CACTCTGACTCCATGGGTTGGTCTTCGGAAGATCAATGTGTCTTACTGGTGCTGGGAGGA 2878
Qy 871 pMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSerAspAlaGl 891
Db 2879 TATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGTGATGCTGG 2938
Qy 891 yPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThrCysIleAs 911
Db 2939 CTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGATTAAAGGCTGCAACCTGCATCAA 2998
Qy 911 nProLeuAsnGlySerValCysGluArgProAlaAsnHisSerAlaLysGlnCysArgTh 931
Db 2999 CCCTCTCAATGGCAGCGTCTGTGAAAGGCCTGCAAACCACAGTGCCAAGCAGTGCCGGAC 3058
Qy 931 rProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTr 951
Db 3059 ACCATGTGCCCTGCGGACAGCGTGTGGCGAGTGCACTAGCAGCAGCTCGGAGTGCATGTG 3118
Qy 951 pCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrValAlaSerProProPheGl 971
Db 3119 GTGCAGTAACATGAAGCAGTGTGTGGACTCCAATGCCTACGTGGCCTCCTTCCCTTTTGG 3178
Qy 971 yGlnCysMetGluTrpTyrThrMetSerThrCysProProGluAsnCysSerGlyTyrCy 991
Db 3179 CCAGTGTATGGAATGGTATACGATGAGCAGCTGCCCACCTGAAAATTGCTCTGGCTACTG 3238
Qy 991 sThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCysThrAspProSerAsnTh 1011
Db 3239 TACCTGCAGCCATTGCTTGGAGCAGCCAGGCTGTGGTTGGTGTACTGATCCTAGCAATAC 3298
Qy 1011 rGlyLysGlyLysCysIleGluGlySerTyrLysGlyProValLysMetProSerGlnAl 1031
Db 3299 TGGGAAAGGAAAATGTATTGAGGGCAGCTATAAAGGACCTGTGAAGATGCCGTCACAGGC 3358
Qy 1031 aProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSe 1051
Db 3359 CTCTGCAGGAAATGTGTATCCACAGCCCCTTCTGAACTCCAGCATGTGTCTAGAGGACAG 3418
Qy 1051 rArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCy 1071
Db 3419 CAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAGTGCAACGGACACAGCAAATG 3478
Qy 1071 sIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrThrGlyLysHisCysGluTh 1091
Db 3479 CATCAACCAGAGTATCTGTGAGAAGTGTGAGGACCTGACCACGGGCAAGCACTGCGAGAC 3538
Qy 1091 rCysIleSerGlyPheTyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLysCy 1111
Db 3539 CTGCATATCTGGCTTCTATGGTGACCCGACTAATGGAGGCAAATGTCAGCCATGCAAGTG 3598
Qy 1111 sAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGl 1131
Db 3599 CAATGGGCACGCATCACTGTGCAACACCAACACCGGCAAGTGCTTCTGTACCACCAAAGG 3658
Qy 1131 yValLysGlyAspGluCysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLe 1151

```
Db       3659 TGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCT 3718

Qy       1151 uArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGl 1171
              |:::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3719 CAAAGGAACATGCTACTATACCCTTCTCATTGACTATCAGTTCACCTTTAGCCTGTCCCA 3778

Qy       1171 nGluAspAspArgTyrTyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnAr 1191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3779 GGAAGACGACCGCTACTACACAGCCATCAACTTTGTGGCTACTCCTGATGAACAAAACAG 3838

Qy       1191 gAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAl 1211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3839 GGATTTGGACATGTTCATCAATGCCTCCAAAAACTTCAACCTCAACATCACCTGGGCCAC 3898

Qy       1211 aSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIl 1231
              ||||||   ||||||||||||   |||||||||:::||||||||||||||||||||||||
Db       3899 CAGCTTCCCAGCCGGAACCCAGACTGGAGAAGAGGTGCCTGTTGTTTCAAAAACCAACAT 3958

Qy       1231 eLysGluTyrLysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIl 1251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3959 CAAGGAATACAAAGATAGCTTCTCTAATGAGAAATTTGATTTTCGCAACCATCCAAACAT 4018

Qy       1251 eThrPhePheValTyrValSerAsnPheThrTrpProIleLysIleGlnIleAlaPheSe 1271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4019 CACTTTCTTTGTTTATGTCAGTAATTTCACTTGGCCCATCAAAATTCAGATTGCCTTCTC 4078

Qy       1271 rGlnHisSerAsnPheMetAspLeuValGlnPhePheValThrPhePheSerCysPheLe 1291
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4079 CCAGCACAGCAACTTCATGGACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCT 4138

Qy       1291 uSerLeuLeuLeuValAlaAlaValValTrpLysIleLysGlnSerCysTrpAlaSerAr 1311
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4139 CTCGCTGCTTCTGGTGGCTGCAGTGGTCTGGAAGATCAAGCAGAGCTGTTGGGCATCCAG 4198

Qy       1311 gArgArgGluGlnLeuLeuArgGluMetGlnGlnMetAlaSerArgProPheAlaSerVa 1331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4199 GCGGAGAGAGCAACTTCTTCGGGAGATGCAACAGATGGCCAGCCGCCCCTTTGCTTCTGT 4258

Qy       1331 lAsnValAlaLeuGluThrAspGluGluProProAspLeuIleGlyGlySerIleLysTh 1351
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4259 AAACGTTGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGAC 4318

Qy       1351 rValProLysProIleAlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerVa 1371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4319 TGTTCCCAAACCCATTGCACTGGAGCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGT 4378

Qy       1371 lPheValArgLeuProArgGlyLeuGlyGlyIleProProProGlyGlnSerGlyLeuAl 1391
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4379 GTTTGTGAGGCTCCCTCGAGGCCTGGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGC 4438

Qy       1391 aValAlaSerAlaLeuValAspIleSerGlnGlnMetProIleValTyrLysGluLysSe 1411
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4439 TGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTC 4498

Qy       1411 rGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnProGlyThrCysIle 1429
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4499 AGGAGCCGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGCCTGGGACCTGCATC 4553

RESULT 7
AF119821
LOCUS       AF119821                4313 bp    mRNA    linear   ROD 16-APR-1999
DEFINITION  Mus musculus attractin (Mgca) mRNA, complete cds.
ACCESSION   AF119821
VERSION     AF119821.1  GI:4585306
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 4313)
  AUTHORS   Gunn,T.M., Miller,K.A., He,L., Hyman,R.W., Davis,R.W., Azarani,A.,
            Schlossman,S.F., Duke-Cohan,J.S. and Barsh,G.S.
  TITLE     The mouse mahogany locus encodes a transmembrane form of human
            attractin
  JOURNAL   Nature 398 (6723), 152-156 (1999)
  MEDLINE   99184160
   PUBMED   10086356
REFERENCE   2  (bases 1 to 4313)
  AUTHORS   Gunn,T.M., Miller,K.A., He,L., Hyman,R.W., Davis,R.W., Azarani,A.,
            Schlossman,S.F., Duke-Cohan,J.S. and Barsh,G.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JAN-1999) HHMI, Stanford, Beckman Center B271A,
            Stanford, CA 94305-5323, USA
FEATURES             Location/Qualifiers
     source          1. .4313
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /chromosome="2"
     gene            1. .4313
                     /gene="Mgca"
                     /note="mahogany candidate"
     CDS             27. .4313
                     /gene="Mgca"
                     /note="mahogany"
                     /codon_start=1
                     /product="attractin"
                     /protein_id="AAD25372.1"
                     /db_xref="GI:4585307"
                     /translation="MVAVAVAASTEARLRGSTTATAAPAGRKGRQHRPCTATGAWRPG
                     PRARLCLPRVLSRALPPPPLLPLLFSLLLLPLPREAEAAAVAAAVSGSAAAEAKECDR
                     PCVNGGRCNPGTGQCVCPTGWVGEQCQHCGGRFRLTGSSGFVTDGPGNYKYKTKCTWL
                     IEGYPNAVLRLRFNHFATECSWDHLYVYDGDSIYAPLIAAFSGLIVPERDGNETAPEV
                     TVTSGYALLHFFSDAAYNLTGFNITYNFDMCPNNCSARGECKSSNSSSAVECECSENW
                     KGESCDIPHCTDNCGFPHRGICNASDTRGCSCFPHWQGPGCSIPVPANQSFWTREEYS
                     DLKLPRASHKAVVNGNIMWVVGGYMFNHSDYSMVLAYDLTSREWLPLNHSVNSVVVRY
                     GHSLALHKDKIYMYGGKIDSTGNVTNELRVFHIHNESWVLLTPKAKDQYAVVGHSAHI
                     VTLASGRVVMLVIFGHCPLYGYISVVQEYDLEKNTWSILHTQGALVQGGYGHSSAYDD
                     RTKALYVHGGYKAFSANKYRLADDLYRYDVDTQMWTILKDSRFFRYLHTAVIVSGTML
                     VFGGNTHNDTSMSHGAKCPSSDFMAYDIACDRWSVLPRPELHHDVNRFGHSAVLYNST
                     MYVFGGFNSLLLSDVLVFTSEQCDAHRSEAACVAAGPGIRCLWDTQSSRCTSWELATE
                     EQAEKLKSECFSKRTLDHDRCDQHTDCYSCTANTNDCHWCNDHCVPVNHSCTEGQISI
                     AKYESCPKDNPMYYCNKKTSCRSCALDQNCQWEPRNQECIALPENICGNGWHLVGNSC
                     LKITTAKENYDNAKLSCRNHNAFLASLTSQKKVELVLKQLRLMQSSQSMSKLTLTPWV
                     GLRKINVSYWCWEDMSPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLNG
                     SVCERPANHSAKQCRTPCALRTACGECTSSSSECMWCSNMKQCVDSNAYVASFPFGQC
                     MEWYTMSSCPPENCSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKMPSQA
                     SAGNVYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSVCEKCEDLTTGKHC
                     ETCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGDECQLCKVENRYQ
                     GNPLKGTCYYTLLIDYQFTFSLSQGDDRYYTAINFVATPDEQNRDFDMFINASKKFNL
                     NITWATSFPAGTQTGEEVPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTWP
                     IKIQIAFSQHSNFMDLVQFFVTFFSCFLSLLLVAAVVWKIKQSCWASRRREQLLREMQ
                     QMASRPFASVNVALETDEEPPDLIGGSIKTVPKPIALEPCFGNKAAVLSVFVRLPRGL
                     GGIPPPGQSGLAVASALVDISQQMPIVYKEKSGAVRNRKQQPPAQPGTCI"
ORIGIN

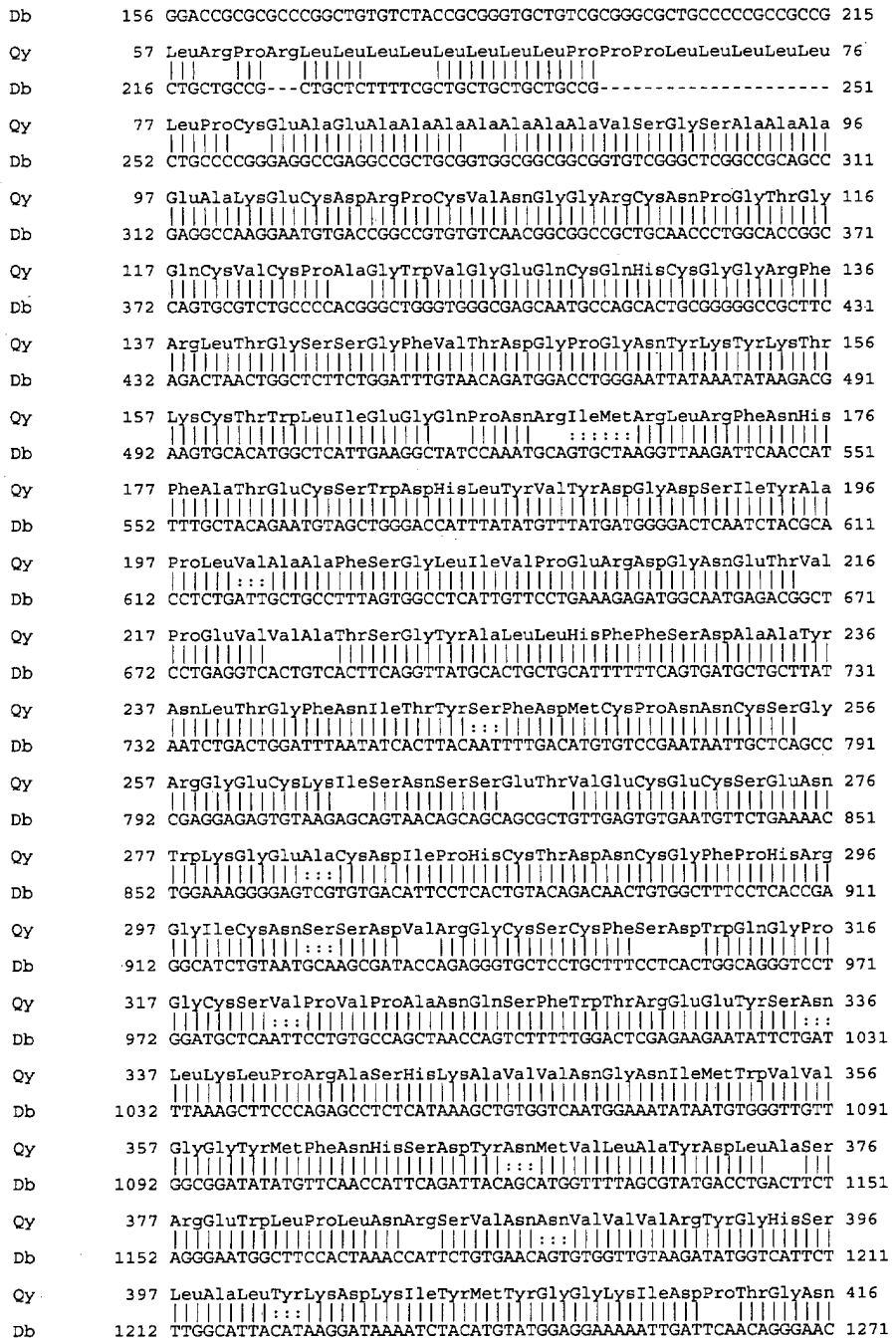
Alignment Scores:
Pred. No.:                0                Length:          4313
Score:                    7302.50          Matches:         1319
Percent Similarity:       94.63%           Conservative:    37
Best Local Similarity:    92.04%           Mismatches:      64
Query Match:              92.26%           Indels:          13
DB:                       10               Gaps:            3

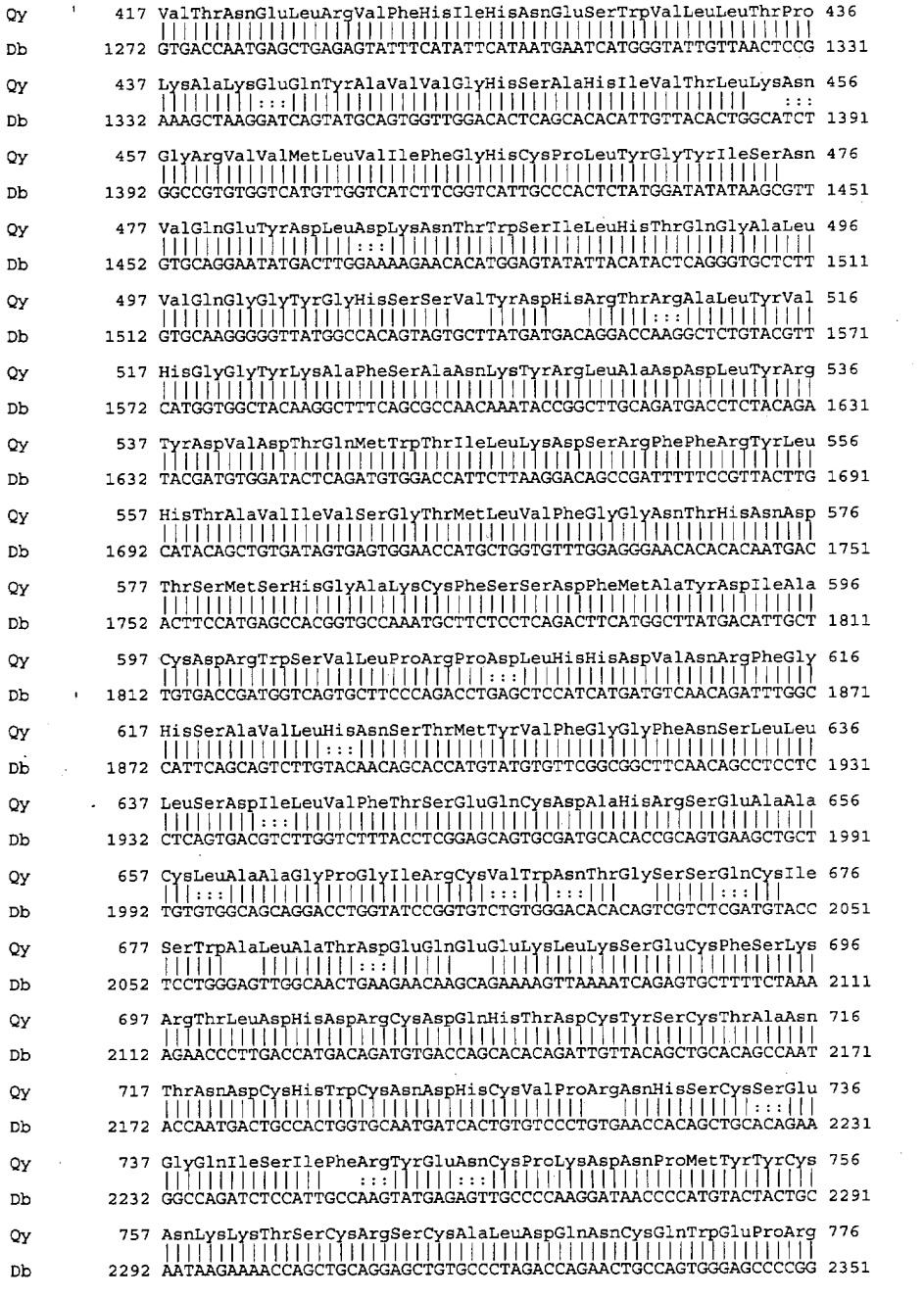
US-09-787-097-12 (1-1429) x AF119821 (1-4313)

Qy          2 ValAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAlaAla 21
              ||||||   ||||||:::|||||||||||||||||||   |||   ||||||||||||
Db         36 GTGGCGGTGGCGGCGTCGACTGAGGCGCGGCTGAGGGGGAGCACGACGGCGACAGCAGCG 95

Qy         22 LeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArgPro 41
                ||||||||   |||      |||                  |||   |||   ||||||
Db         96 CCTGCGGGCAGGAAGGGCAGGCAGCACCGACCCTGCACCGCGACAGGGGCCTGGAGGCCG 155

Qy         42 GlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSer----------------ProPro 56
                |||      |||   |||   |||||||||:::||||||                ||||||
```


Db 156 GGACCGCGCGCCCGGCTGTGTCTACCGCGGGTGCTGTCGCGGGCGCTGCCCCCGCCGCCG 215
Qy 57 LeuArgProArgLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeu 76
Db 216 CTGCTGCCG---CTGCTCTTTTCGCTGCTGCTGCTGCCG---------- 251
Qy 77 LeuProCysGluAlaGluAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAla 96
Db 252 CTGCCCCGGGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCGGGCTCGGCCGCAGCC 311
Qy 97 GluAlaLysGluCysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGly 116
Db 312 GAGGCCAAGGAATGTGACCGGCCGTGTGTCAACGGCGGCCGCTGCAACCCTGGCACCGGC 371
Qy 117 GlnCysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPhe 136
Db 372 CAGTGCGTCTGCCCCACGGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTC 431
Qy 137 ArgLeuThrGlySerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThr 156
Db 432 AGACTAACTGGCTCTTCTGGATTTGTAACAGATGGACCTGGGAATTATAAATATAAGACG 491
Qy 157 LysCysThrTrpLeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHis 176
Db 492 AAGTGCACATGGCTCATTGAAGGCTATCCAAATGCAGTGCTAAGGTTAAGATTCAACCAT 551
Qy 177 PheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAla 196
Db 552 TTTGCTACAGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATCTACGCA 611
Qy 197 ProLeuValAlaAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrVal 216
Db 612 CCTCTGATTGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCT 671
Qy 217 ProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyr 236
Db 672 CCTGAGGTCACTGTCACTTCAGGTTATGCACTGCTGCATTTTTTCAGTGATGCTGCTTAT 731
Qy 237 AsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGly 256
Db 732 AATCTGACTGGATTTAATATCACTTACAATTTTGACATGTGTCCGAATAATTGCTCAGCC 791
Qy 257 ArgGlyGluCysLysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsn 276
Db 792 CGAGGAGAGTGTAAGAGCAGTAACAGCAGCAGCGCTGTTGAGTGTGAATGTTCTGAAAAC 851
Qy 277 TrpLysGlyGluAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArg 296
Db 852 TGGAAAGGGGAGTCGTGTGACATTCCTCACTGTACAGACAACTGTGGCTTTCCTCACCGA 911
Qy 297 GlyIleCysAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyPro 316
Db 912 GGCATCTGTAATGCAAGCGATACCAGAGGGTGCTCCTGCTTTCCTCACTGGCAGGGTCCT 971
Qy 317 GlyCysSerValProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsn 336
Db 972 GGATGCTCAATTCCTGTGCCAGCTAACCAGTCTTTTTGGACTCGAGAAGAATATTCTGAT 1031
Qy 337 LeuLysLeuProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValVal 356
Db 1032 TTAAAGCTTCCCAGAGCCTCTCATAAAGCTGTGGTCAATGGAAATATAATGTGGGTTGTT 1091
Qy 357 GlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSer 376
Db 1092 GGCGGATATATGTTCAACCATTCAGATTACAGCATGGTTTTAGCGTATGACCTGACTTCT 1151
Qy 377 ArgGluTrpLeuProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSer 396
Db 1152 AGGGAATGGCTTCCACTAAACCATTCTGTGAACAGTGTGGTTGTAAGATATGGTCATTCT 1211
Qy 397 LeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsn 416
Db 1212 TTGGCATTACATAAGGATAAAATCTACATGTATGGAGGAAAAATTGATTCAACAGGGAAC 1271


Qy 417 ValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrPro 436
Db 1272 GTGACCAATGAGCTGAGAGTATTTCATATTCATAATGAATCATGGGTATTGTTAACTCCG 1331
Qy 437 LysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsn 456
Db 1332 AAAGCTAAGGATCAGTATGCAGTGGTTGGACACTCAGCACACATTGTTACACTGGCATCT 1391
Qy 457 GlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsn 476
Db 1392 GGCCGTGTGGTCATGTTGGTCATCTTCGGTCATTGCCCACTCTATGGATATATAAGCGTT 1451
Qy 477 ValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeu 496
Db 1452 GTGCAGGAATATGACTTGGAAAAGAACACATGGAGTATATTACATACTCAGGGTGCTCTT 1511
Qy 497 ValGlnGlyGlyTyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrVal 516
Db 1512 GTGCAAGGGGGTTATGGCCACAGTAGTGCTTATGATGACAGGACCAAGGCTCTGTACGTT 1571
Qy 517 HisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArg 536
Db 1572 CATGGTGGCTACAAGGCTTTCAGCGCCAACAAATACCGGCTTGCAGATGACCTCTACAGA 1631
Qy 537 TyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeu 556
Db 1632 TACGATGTGGATACTCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTG 1691
Qy 557 HisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAsp 576
Db 1692 CATACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGAGGGAACACACACAATGAC 1751
Qy 577 ThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAla 596
Db 1752 ACTTCCATGAGCCACGGTGCCAAATGCTTCTCCTCAGACTTCATGGCTTATGACATTGCT 1811
Qy 597 CysAspArgTrpSerValLeuProArgProAspLeuHisHisAspValAsnArgPheGly 616
Db 1812 TGTGACCGATGGTCAGTGCTTCCCAGACCTGAGCTCCATCATGATGTCAACAGATTTGGC 1871
Qy 617 HisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeu 636
Db 1872 CATTCAGCAGTCTTGTACAACAGCACCATGTATGTGTTCGGCGGCTTCAACAGCCTCCTC 1931
Qy 637 LeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAla 656
Db 1932 CTCAGTGACGTCTTGGTCTTTACCTCGGAGCAGTGCGATGCACACCGCAGTGAAGCTGCT 1991
Qy 657 CysLeuAlaAlaGlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIle 676
Db 1992 TGTGTGGCAGCAGGACCTGGTATCCGGTGTCTGTGGGACACACAGTCGTCTCGATGTACC 2051
Qy 677 SerTrpAlaLeuAlaThrAspGluGlnGluGluLysLeuLysSerGluCysPheSerLys 696
Db 2052 TCCTGGGAGTTGGCAACTGAAGAACAAGCAGAAAAGTTAAAATCAGAGTGCTTTTCTAAA 2111
Qy 697 ArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsn 716
Db 2112 AGAACCCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAAT 2171
Qy 717 ThrAsnAspCysHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGlu 736
Db 2172 ACCAATGACTGCCACTGGTGCAATGATCACTGTGTCCCTGTGAACCACAGCTGCACAGAA 2231
Qy 737 GlyGlnIleSerIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCys 756
Db 2232 GGCCAGATCTCCATTGCCAAGTATGAGAGTTGCCCCAAGGATAACCCCATGTACTACTGC 2291
Qy 757 AsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArg 776
Db 2292 AATAAGAAAACCAGCTGCAGGAGCTGTGCCCTAGACCAGAACTGCCAGTGGGAGCCCCGG 2351

```
Qy       777 AsnGlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGly 796
Db      2352 AATCAAGAGTGCATCGCCCTGCCGGAAAATATCTGTGGCAATGGCTGGCATTTGGTTGGA 2411

Qy       797 AsnSerCysLeuLysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCys 816
Db      2412 AACTCGTGTCTGAAAATCACTACTGCTAAGGAGAATTATGACAATGCTAAATTGTCCTGT 2471

Qy       817 ArgAsnHisAsnAlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeu 836
Db      2472 AGGAACCACAATGCCTTTTTGGCTTCCCTCACATCCCAGAAGAAGGTGGAGCTTGTCCTT 2531

Qy       837 LysGlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrp 856
Db      2532 AAGCAGCTTCGATTAATGCAATCATCTCAAAGTATGTCCAAGCTCACTCTGACTCCATGG 2591

Qy       857 ValGlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThr 876
Db      2592 GTTGGTCTTCGGAAGATCAATGTGTCTTACTGGTGCTGGGAGGATATGTCTCCATTCACA 2651

Qy       877 AsnSerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeu 896
Db      2652 AATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGTGATGCTGGCTTCTGTGGGATCTTG 2711

Qy       897 SerGluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySer 916
Db      2712 TCAGAGCCTAGTACTCGGGGATTAAAGGCTGCAACCTGCATCAACCCTCTCAATGGCAGC 2771

Qy       917 ValCysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArg 936
Db      2772 GTCTGTGAAAGGCCTGCAAACCACAGTGCCAAGCAGTGCCGGACACCATGTGCCCTGCGG 2831

Qy       937 ThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLys 956
Db      2832 ACAGCGTGTGGCGAGTGCACTAGCAGCAGCTCGGAGTGCATGTGGTGCAGTAACATGAAG 2891

Qy       957 GlnCysValAspSerAsnAlaTyrValAlaSerProProPheGlyGlnCysMetGluTrp 976
Db      2892 CAGTGTGTGGACTCCAATGCCTACGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGG 2951

Qy       977 TyrThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCys 996
Db      2952 TATACGATGAGCAGCTGCCCACCTGAAAATTGCTCTGGCTACTGTACCTGCAGCCATTGC 3011

Qy       997 LeuGluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCys 1016
Db      3012 TTGGAGCAGCCAGGCTGTGGTTGGTGTACTGATCCTAGCAATACTGGGAAAGGAAAATGT 3071

Qy      1017 IleGluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPhe 1036
Db      3072 ATTGAGGGCAGCTATAAAGGACCTGTGAAGATGCCGTCACAGGCCTCTGCAGGAAATGTG 3131

Qy      1037 TyrProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSer 1056
Db      3132 TATCCACAGCCCCTTCTGAACTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCT 3191

Qy      1057 PheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIle 1076
Db      3192 TTCATTCACTGTCCAGCTTGCCAGTGCAACGGACACAGCAAATGCATCAACCAGAGTGTC 3251

Qy      1077 CysGluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPhe 1096
Db      3252 TGTGAGAAGTGTGAGGACCTGACCACGGGCAAGCACTGCGAGACCTGCATATCTGGCTTC 3311

Qy      1097 TyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSer 1116
Db      3312 TATGGTGACCCGACTAATGGAGGCAAATGTCAGCCATGCAAGTGCAATGGGCACGCATCA 3371

Qy      1117 LeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGlu 1136
Db      3372 CTGTGCAACACCAACACCGGCAAGTGCTTCTGTACCACCAAAGGTGTCAAGGGGGACGAG 3431

Qy      1137 CysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyr 1156
Db      3432 TGCCAGCTATGTAAGGTAGAAAATCGATACCAGGGGAACCCTCTCAAAGGAACATGCTAC 3491

Qy      1157 TyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyr 1176
Db      3492 TATACCCTTCTCATTGACTATCAGTTCACCTTTAGCCTGTCCCAGGGAGACGACCGCTAC 3551

Qy      1177 TyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPhe 1196
Db      3552 TACACAGCCATCAACTTTGTGGCTACTCCTGATGAACAAAACAGGGATTTTGACATGTTC 3611

Qy      1197 IleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGly 1216
Db      3612 ATCAATGCCTCCAAAAAATTCAACCTCAACATCACCTGGGCCACCAGCTTCCCAGCCGGA 3671

Qy      1217 ThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAsp 1236
Db      3672 ACCCAGACTGGAGAAGAGGTGCCTGTTGTTTCAAAAACCAACATCAAGGAATACAAAGAT 3731

Qy      1237 SerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyr 1256
Db      3732 AGCTTCTCTAATGAGAAATTTGATTTTCGCAACCATCCAAACATCACTTTCTTTGTTTAT 3791

Qy      1257 ValSerAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPhe 1276
Db      3792 GTCAGTAATTTCACTTGGCCCATCAAAATTCAGATTGCCTTCTCCCAGCACAGCAACTTT 3851

Qy      1277 MetAspLeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuLeuVal 1296
Db      3852 ATGGACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCGCTGCTTCTGGTG 3911

Qy      1297 AlaAlaValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeu 1316
Db      3912 GCTGCAGTGGTCTGGAAGATCAAGCAGAGCTGTTGGGCATCCAGGCGGAGAGAGCAACTT 3971

Qy      1317 LeuArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGlu 1336
Db      3972 CTTCGGGAGATGCAACAGATGGCCAGCCGCCCCTTTGCTTCTGTAAACGTTGCCTTGGAA 4031

Qy      1337 ThrAspGluGluProProAspLeuIleGlyGlySerIleLysThrValProLysProIle 1356
Db      4032 ACAGATGAAGAACCTCCTGATCTCATTGGGGGAAGTATAAAGACCGTTCCTAAGCCCATT 4091

Qy      1357 AlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuPro 1376
Db      4092 GCCCTGGAGCCCTGCTTTGGTAACAAAGCCGCAGTCCTCTCTGTATTCGTGAGGCTCCCT 4151

Qy      1377 ArgGlyLeuGlyGlyIleProProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeu 1396
Db      4152 CGAGGACTGGGAGGAATCCCTCCTCCTGGTCAGTCAGGTCTCGCTGTGGCCAGTGCCCTG 4211

Qy      1397 ValAspIleSerGlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAsn 1416
Db      4212 GTGGACATTTCTCAGCAGATGCCAATAGTGTACAAGGAGAAGTCAGGAGCTGTAAGAAAC 4271

Qy      1417 ArgLysGlnGlnProProAlaGlnProGlyThrCysIle 1429
Db      4272 CGGAAGCAGCAGCCGCCTGCACAGCCTGGAACCTGCATT 4310
```

```
RESULT 8
AF106861
LOCUS       AF106861               3819 bp    mRNA    linear   PRI 06-JAN-1999
DEFINITION  Homo sapiens attractin-2 (ATRN) mRNA, complete cds.
ACCESSION   AF106861
VERSION     AF106861.1  GI:4093195
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3819)
  AUTHORS   Duke-Cohan,J.S., Gu,J., Freeman,G.J. and Schlossman,S.F.
  TITLE     Cloning of cDNA for attractin-2, identical with that of attractin
```

```
            except for a GC-rich 222 bp 5' insertion
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3819)
  AUTHORS   Duke-Cohan,J.S., Gu,J., Freeman,G.J. and Schlossman,S.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-NOV-1998) Division of Tumor Immunology, Dana Farber
            Cancer Institute, 44 Binney Street, Boston, MA 02115, USA
FEATURES             Location/Qualifiers
     source          1. .3819
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     gene            1. .3819
                     /gene="ATRN"
     CDS             1. .3819
                     /gene="ATRN"
                     /note="similar to Homo sapiens attractin in GenBank
                     Accession Number AF034957"
                     /codon_start=1
                     /product="attractin-2"
                     /protein_id="AAD03057.1"
                     /db_xref="GI:4093196"
                     /translation="MVAAAAATEARLRRRTAATAALAGRSGGPHWDWDVTRAGRPGLG
                     AGLRLPRLLSPPLRPRLLLLLLLLPPPLLLLLLPCEAEAAAAAAAVSGSAAAEAKECD
                     RPCVNGGRCNPGTGQCVCPAGWVGEQCQHCGGRFRLTGSSGFVTDGPGNYKYKTKCTW
                     LIEGQPNRIMRLRFNHFATECSWDHLYVYDGDSIYAPLVAAFSGLIVPERDGNETVPE
                     VVATSGYALLHFFSDAAYNLTGFNITYSFDMCPNNCSGRGECKISNSSETVECECSEN
                     WKGEACDIPHCTDNCGFPHRGICNSSDVRGCSCFSDWQGPGCSVPVPANQSFWTREEY
                     SNLKLPRASHKAVVNGNIMWVVGGYMFNHSDYNMVLAYDLASREWLPLNRSVNNVVVR
                     YGHSLALYKDKIYMYGGKIDPTGNVTNELRVFHIHNESWVLLTPKAKEQYAVVGHSAH
                     IVTLKNGRVVMLVIFGHCPLYGYISNVQEYDLDKNTWSILHTQGALVQGGYGHSSVYD
                     HRTRALYVHGGYKAFSANKYRLADDLYRYDVDTQMWTILKDSRFFRYLHTAVIVSGTM
                     LVFGGNTHNDTSMSHGAKCFSSDFMAYDIACDRWSVLPRPDLHHDVNRFGHSAVLHNS
                     TMYVFGGFNSLLLSDILVFTSEQCDAHRSEAACLAAGPGIRCVWNTGSSQCISWALAT
                     DEQEEKLKSECFSKRTLDHDRCDQHTDCYSCTANTNDCHWCNDHCVPRNHSCSEGQIS
                     IFRYENCPKDNPMYYCNKKTSCRSCALDQNCQWEPRNQECIALPENICGIGWHLVGNS
                     CLKITTAKENYDNAKLFCRNHNALLASLTTQKKVEFVLKQLRIMQSSQSMSKLTLTPW
                     VGLRKINVSYWCWEDMSPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLN
                     GSVCERPANHSAKQCRTPCALRTACGDCTSGSSECMWCSNMKQCVDSNAYVASFPFGQ
                     CMEWYTMSTCPPENCSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKMPSQ
                     APTGNFYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCENLTTGKH
                     CETCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGDECQLCEVENRY
                     QGNPLRGTCYYTLLIDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASKNFN
                     LNITWAASFSAGTQAGEEMPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTW
                     PIKIQVQTEQ"
ORIGIN

Alignment Scores:
Pred. No.:                0          Length:         3819
Score:                    7089.00    Matches:        1268
Percent Similarity:       99.76%     Conservative:   1
Best Local Similarity:    99.69%     Mismatches:     3
Query Match:              89.56%     Indels:         0
DB:                       9          Gaps:           0

US-09-787-097-12 (1-1429) x AF106861 (1-3819)

Qy          1 MetValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 60

Qy         21 AlaLeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArg 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGGGACTGGGACGTGACCAGGGCTGGGAGG 120

Qy         41 ProGlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSerProProLeuArgProArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGGGGCTGGGGGCCGGGCTGCGCCTCCCGCGGCTGCTGTCTCCACCGCTGCGGCCACGG 180

Qy         61 LeuLeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeuLeuProCysGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGCTGCTGCTGCTGTTGTTGCTCCCGCCGCCGCTGTTGCTGCTGCTGCTGCCCTGTGAG 240

Qy         81 AlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCGAGGCCGCGGCGGCGGCGGCGGCGGTGTCGGGCTCAGCCGCAGCCGAGGCCAAGGAA 300

Qy        101 CysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCys 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGTGACCGGCCCTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGC 360

Qy        121 ProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGly 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGA 420

Qy        141 SerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrp 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGG 480

Qy        161 LeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAG 540

Qy        181 CysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAla 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCT 600

Qy        201 AlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGluValVal 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTT 660

Qy        221 AlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGly 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGA 720

Qy        241 PheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCys 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGT 780

Qy        261 LysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsnTrpLysGlyGlu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAGATCAGTAATAGCAGCGAAACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAA 840

Qy        281 AlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsn 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAAT 900

Qy        301 SerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerVal 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTT 960

Qy        321 ProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuPro 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCC 1020

Qy        341 ArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMet 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATG 1080

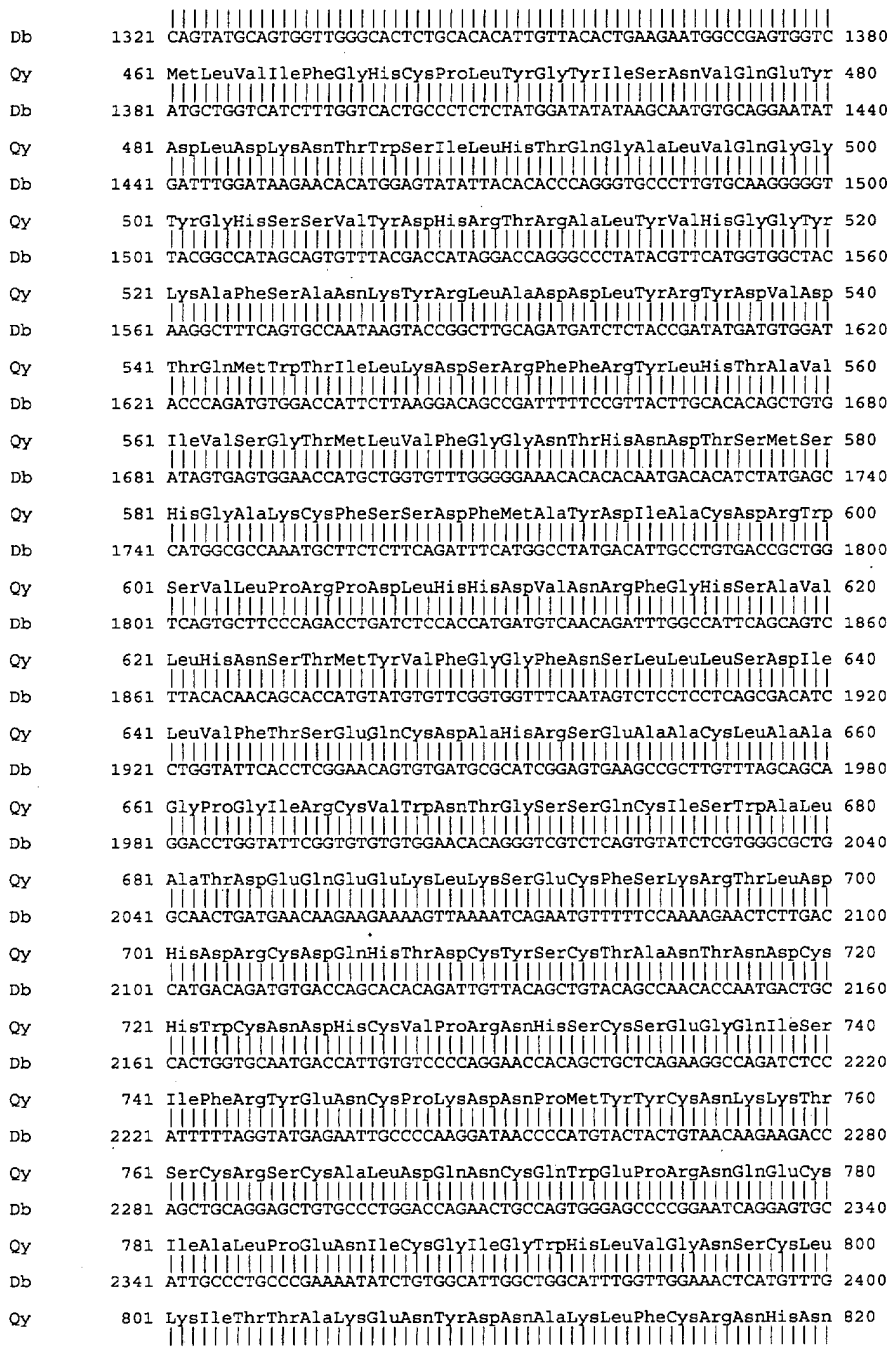
Qy        361 PheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeu 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTT 1140

Qy        381 ProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATAC 1200

Qy        401 LysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsnValThrAsnGlu 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AAGGATAAAATTTACATGTATGGAGGAAAAATTGATCCAACTGGGAATGTGACCAATGAG 1260

Qy        421 LeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGlu 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGGAG 1320

Qy        441 GlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValVal 460
```


Db 1321 CAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTC 1380
Qy 461 MetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyr 480
Db 1381 ATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATAT 1440
Qy 481 AspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGly 500
Db 1441 GATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGT 1500
Qy 501 TyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyr 520
Db 1501 TACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTAC 1560
Qy 521 LysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAsp 540
Db 1561 AAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGAT 1620
Qy 541 ThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaVal 560
Db 1621 ACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTG 1680
Qy 561 IleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSer 580
Db 1681 ATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGC 1740
Qy 581 HisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrp 600
Db 1741 CATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGG 1800
Qy 601 SerValLeuProArgProAspLeuHisHisAspValAsnArgPheGlyHisSerAlaVal 620
Db 1801 TCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTC 1860
Qy 621 LeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIle 640
Db 1861 TTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATC 1920
Qy 641 LeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAla 660
Db 1921 CTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCA 1980
Qy 661 GlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeu 680
Db 1981 GGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTG 2040
Qy 681 AlaThrAspGluGlnGluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAsp 700
Db 2041 GCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGAC 2100
Qy 701 HisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCys 720
Db 2101 CATGACAGATGTGACCAGCACACAGATTGTTACAGCTGTACAGCCAACACCAATGACTGC 2160
Qy 721 HisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSer 740
Db 2161 CACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCC 2220
Qy 741 IlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThr 760
Db 2221 ATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCCATGTACTACTGTAACAAGAAGACC 2280
Qy 761 SerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCys 780
Db 2281 AGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGC 2340
Qy 781 IleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeu 800
Db 2341 ATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTG 2400
Qy 801 LysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsn 820
Db 2401 AAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAAT 2460
Qy 821 AlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArg 840
Db 2461 GCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGA 2520
Qy 841 IleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArg 860
Db 2521 ATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGG 2580
Qy 861 LysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeu 880
Db 2581 AAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTA 2640
Qy 881 GlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSer 900
Db 2641 CAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGT 2700
Qy 901 ThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArg 920
Db 2701 ACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGG 2760
Qy 921 ProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGly 940
Db 2761 CCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGA 2820
Qy 941 AspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAsp 960
Db 2821 GATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGAC 2880
Qy 961 SerAsnAlaTyrValAlaSerProProPheGlyGlnCysMetGluTrpTyrThrMetSer 980
Db 2881 TCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGC 2940
Qy 981 ThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPro 1000
Db 2941 ACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCA 3000
Qy 1001 GlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySer 1020
Db 3001 GGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCC 3060
Qy 1021 TyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnPro 1040
Db 3061 TATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCC 3120
Qy 1041 LeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCys 1060
Db 3121 CTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGT 3180
Qy 1061 ProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCys 1080
Db 3181 CCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGT 3240
Qy 1081 GluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspPro 1100
Db 3241 GAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCC 3300
Qy 1101 ThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThr 1120
Db 3301 ACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACC 3360
Qy 1121 AsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCys 1140
Db 3361 AACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGT 3420
Qy 1141 GluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeu 1160
Db 3421 GAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTT 3480
Qy 1161 IleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIle 1180
Db 3481 ATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATC 3540

```
Qy       1181 AsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSer 1200
Db       3541 AATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCC 3600

Qy       1201 LysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGly 1220
Db       3601 AAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGA 3660

Qy       1221 GluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsn 1240
Db       3661 GAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAAT 3720

Qy       1241 GluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPhe 1260
Db       3721 GAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTC 3780

Qy       1261 ThrTrpProIleLysIleGlnIleAlaPheSerGln 1272
Db       3781 ACCTGGCCCATCAAAATTCAGGTGCAAACTGAACAA 3816

RESULT 9
AR164817
LOCUS       AR164817               4072 bp    DNA     linear   PAT 17-OCT-2001
DEFINITION  Sequence 16 from patent US 6274339.
ACCESSION   AR164817
VERSION     AR164817.1  GI:16238027
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 4072)
  AUTHORS   Moore,K. and Nagle,D.Lynn.
  TITLE     Methods and compositions for the diagnosis and treatment of body
            weight disorders, including obesity
  JOURNAL   Patent: US 6274339-A 16 14-AUG-2001;
FEATURES             Location/Qualifiers
     source          1. .4072
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:                0          Length:          4072
Score:                    6870.50    Matches:         1240
Percent Similarity:       97.41%     Conservative:    3
Best Local Similarity:    97.18%     Mismatches:      20
Query Match:              86.80%     Indels:          13
DB:                       6          Gaps:            3

US-09-787-097-12 (1-1429) x AR164817 (1-4072)

Qy          2 ValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAlaAla 21
Db         10 GTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCG 69

Qy         22 LeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArgPro 41
Db         70 CTCGCGGGCAGGAGCGGCGGGCCGCACCGACCCTGCACCGCGACAGGGGCCTGGAGGCCG 129

Qy         42 GlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSer----------ProPro 56
Db        130 GGACCGCGCGCCCGGCTGTGTCTCCCGCGGGTGCTGTCGCGGGCGCTGCCCCCGCCGCCG 189

Qy         57 LeuArgProArgLeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeu 76
Db        190 CTGCTGCCG---CTGCTCTTTTCGCTGCTGCTGCTGCCG---------- 225

Qy         77 LeuProCysGluAlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAla 96
Db        226 CTGCCCCGGGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCC 285

Qy         97 GluAlaLysGluCysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGly 116
Db        286 GAGGCCAAGGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGC 345

Qy        117 GlnCysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPhe 136
Db        346 CAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTC 405

Qy        137 ArgLeuThrGlySerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThr 156
Db        406 AGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACG 465

Qy        157 LysCysThrTrpLeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHis 176
Db        466 AAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCAT 525

Qy        177 PheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAla 196
Db        526 TTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCA 585

Qy        197 ProLeuValAlaAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrVal 216
Db        586 CCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTC 645

Qy        217 ProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyr 236
Db        646 CCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTAT 705

Qy        237 AsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGly 256
Db        706 AATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGC 765

Qy        257 ArgGlyGluCysLysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsn 276
Db        766 CGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAAC 825

Qy        277 TrpLysGlyGluAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArg 296
Db        826 TGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGA 885

Qy        297 GlyIleCysAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyPro 316
Db        886 GGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCT 945

Qy        317 GlyCysSerValProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsn 336
Db        946 GGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAAC 1005

Qy        337 LeuLysLeuProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValVal 356
Db       1006 TTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTT 1065

Qy        357 GlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSer 376
Db       1066 GGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCT 1125

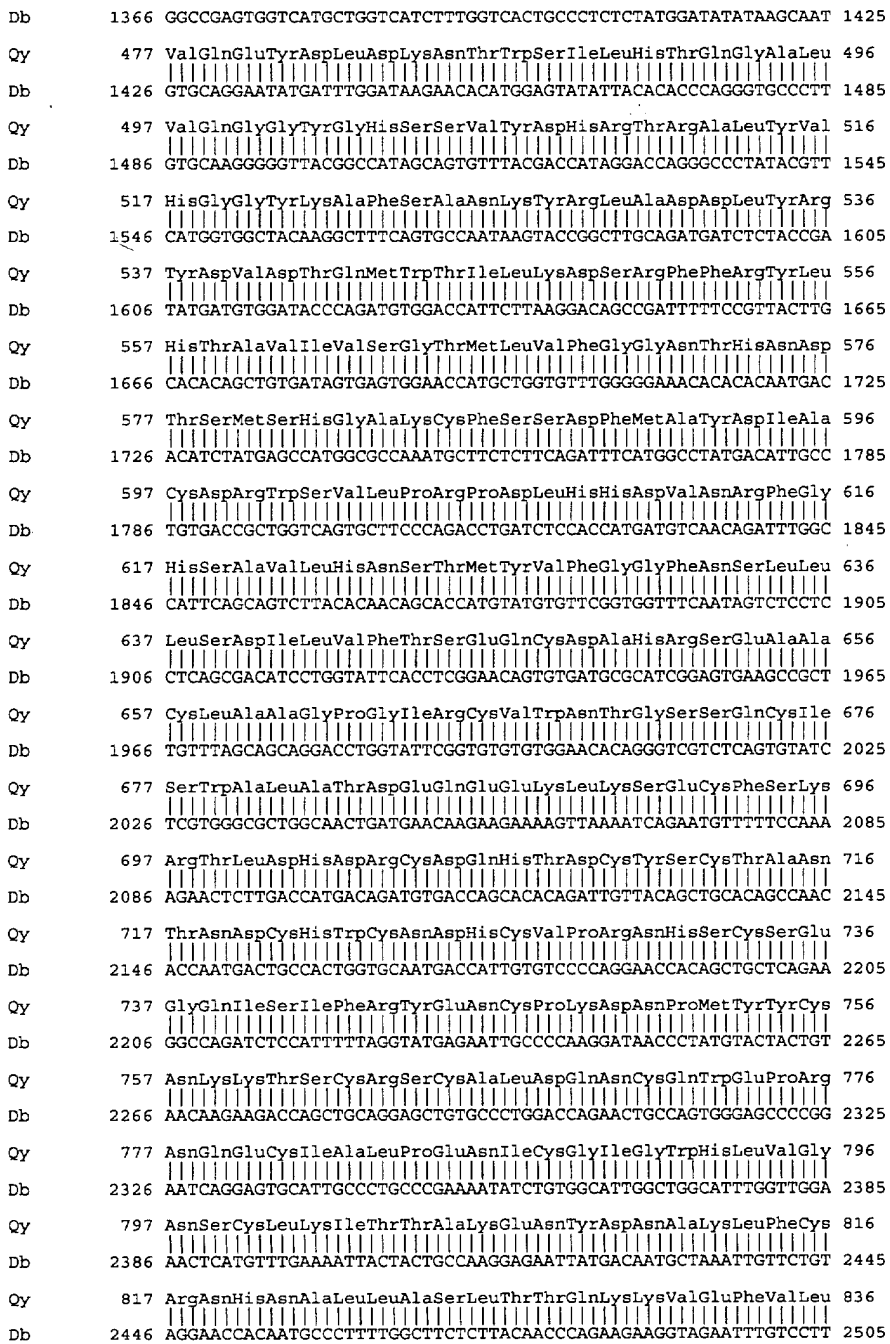
Qy        377 ArgGluTrpLeuProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSer 396
Db       1126 AGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCT 1185

Qy        397 LeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsn 416
Db       1186 TTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAAT 1245

Qy        417 ValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrPro 436
Db       1246 GTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCT 1305

Qy        437 LysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsn 456
Db       1306 AAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAAT 1365

Qy        457 GlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsn 476
```


Db 1366 GGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAAT 1425
Qy 477 ValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeu 496
Db 1426 GTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTT 1485
Qy 497 ValGlnGlyGlyTyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrVal 516
Db 1486 GTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTT 1545
Qy 517 HisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArg 536
Db 1546 CATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGA 1605
Qy 537 TyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeu 556
Db 1606 TATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTG 1665
Qy 557 HisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAsp 576
Db 1666 CACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGAC 1725
Qy 577 ThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAla 596
Db 1726 ACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCC 1785
Qy 597 CysAspArgTrpSerValLeuProArgProAspLeuHisHisAspValAsnArgPheGly 616
Db 1786 TGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGC 1845
Qy 617 HisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeu 636
Db 1846 CATTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTC 1905
Qy 637 LeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAla 656
Db 1906 CTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCT 1965
Qy 657 CysLeuAlaAlaGlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIle 676
Db 1966 TGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATC 2025
Qy 677 SerTrpAlaLeuAlaThrAspGluGlnGluGluLysLeuLysSerGluCysPheSerLys 696
Db 2026 TCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAA 2085
Qy 697 ArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsn 716
Db 2086 AGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAAC 2145
Qy 717 ThrAsnAspCysHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGlu 736
Db 2146 ACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAA 2205
Qy 737 GlyGlnIleSerIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCys 756
Db 2206 GGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGT 2265
Qy 757 AsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArg 776
Db 2266 AACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGG 2325
Qy 777 AsnGlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGly 796
Db 2326 AATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGA 2385
Qy 797 AsnSerCysLeuLysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCys 816
Db 2386 AACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGT 2445
Qy 817 ArgAsnHisAsnAlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeu 836
Db 2446 AGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTT 2505
Qy 837 LysGlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrp 856
Db 2506 AAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGG 2565
Qy 857 ValGlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThr 876
Db 2566 GTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACA 2625
Qy 877 AsnSerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeu 896
Db 2626 AATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTA 2685
Qy 897 SerGluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySer 916
Db 2686 TCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGT 2745
Qy 917 ValCysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArg 936
Db 2746 GTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGG 2805
Qy 937 ThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLys 956
Db 2806 ACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAG 2865
Qy 957 GlnCysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrp 976
Db 2866 CAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGG 2925
Qy 977 TyrThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCys 996
Db 2926 TATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGC 2985
Qy 997 LeuGluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCys 1016
Db 2986 TTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGC 3045
Qy 1017 IleGluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPhe 1036
Db 3046 ATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTC 3105
Qy 1037 TyrProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSer 1056
Db 3106 TATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCT 3165
Qy 1057 PheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIle 1076
Db 3166 TTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATC 3225
Qy 1077 CysGluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPhe 1096
Db 3226 TGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTC 3285
Qy 1097 TyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSer 1116
Db 3286 TACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCT 3345
Qy 1117 LeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGlu 1136
Db 3346 CTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAG 3405
Qy 1137 CysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyr 1156
Db 3406 TGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTAT 3465
Qy 1157 TyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyr 1176
Db 3466 TATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTAT 3525
Qy 1177 TyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPhe 1196
Db 3526 TACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTC 3585

```
Qy       1197 IleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGly 1216
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3586 ATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGA 3645

Qy       1217 ThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAsp 1236
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3646 ACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGAT 3705

Qy       1237 SerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyr 1256
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3706 AGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTAT 3765

Qy       1257 ValSerAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGln 1272
             ||||||||||||||||||||||||||||||:::          |||
Db       3766 GTCAGTAATTTCACCTGGCCCATCAAAATTCAGGTGCAAACTGAACAA 3813

RESULT 10
AF034957
LOCUS       AF034957               3597 bp    mRNA    linear   PRI 30-SEP-1998
DEFINITION  Homo sapiens secreted T cell activation protein Attractin
            (attractin) mRNA, complete cds.
ACCESSION   AF034957
VERSION     AF034957.1  GI:3676346
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3597)
  AUTHORS   Duke-Cohan,J.S., Gu,J., McLaughlin,D.F., Xu,Y., Freeman,G.J. and
            Schlossman,S.F.
  TITLE     Attractin (DPPT-L), a member of the CUB family of cell adhesion and
            guidance proteins, is secreted by activated human T lymphocytes and
            modulates immune cell interactions
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 95 (19), 11336-11341 (1998)
  MEDLINE   98409658
   PUBMED   9736737
REFERENCE   2  (bases 1 to 3597)
  AUTHORS   Duke-Cohan,J.S., Gu,J., Ao,Z., McLaughlin,D.F., Freeman,G.J. and
            Schlossman,S.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-NOV-1997) Division of Cancer, Immunology and AIDS,
            Dana Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
            USA
FEATURES             Location/Qualifiers
     source          1. .3597
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     gene            1. .3597
                     /gene="attractin"
     CDS             1. .3597
                     /gene="attractin"
                     /note="secreted T cell activation protein; Method:
                     conceptual translation with partial peptide sequencing;
                     DPPT-L"
                     /codon_start=1
                     /product="Attractin"
                     /protein_id="AAC61902.1"
                     /db_xref="GI:3676347"
                     /translation="MVAAAAATEARLRRRTAATAALAGRSGGPHCVNGGRCNPGTGQC
                     VCPAGWVGEQCQHCGGRFRLTGSSGFVTDGPGNYKYKTKCTWLIEGQPNRIMRLRFNH
                     FATECSWDHLYVYDGDSIYAPLVAAFSGLIVPERDGNETVPEVVATSGYALLHFFSDA
                     AYNLTGFNITYSFDMCPNNCSGRGECKISNSSDTVECECSENWKGEACDIPHCTDNCG
                     FPHRGICNSSDVRGCSCFSDWQGPGCSVPVPANQSFWTREEYSNLKLPRASHKAVVNG
                     NIMWVVGGYMFNHSDYNMVLAYDLASREWLPLNRSVNNVVVRYGHSLALYKDKIYMYG
                     GKIDSTGNVTNELRVFHIHNESWVLLTPKAKEQYAVVGHSAHIVTLKNGRVVMLVIFG
                     HCPLYGYISNVQEYDLDKNTWSILHTQGALVQGGYGHSSVYDHRTRALYVHGGYKAFS
                     ANKYRLADDLYRYDVDTQMWTILKDSRFFRYLHTAVIVSGTMLVFGGNTHNDTSMSHG
                     AKCFSSDFMAYDIACDRWSVLPRPDSTMMSTDLAIPAVLHNSTMYVFGGFNSLLLSDI
                     LVFTSEQCDAHRSEAACLAAGPGIRCVWNTGSSQCISWALATDEQEEKLKSECFSKRT
                     LDHDRCDQHTDCYSCTANTNDCHWCNDHCVPRNHSCSEGQISIFRYENCPKDNPMYYC
                     NKKTSCRSCALDQNCQWEPRNQECIALPENICGIGWHLVGNSCLKITTAKENYDNAKL
                     FCRNHNALLASLTTQKKVEPVLKQLRIMQSSQSMSKLTLTPWVGLRKINVSYWCWEDM
                     SPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLNGSVCERPANHSAKQCR
                     TPCALRTACGDCTSGSSECMWCSNMKQCVDSNAYVASFPFGQCMEWYTMSTCPPENCS
                     GYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKMPSQAPTGNFYPQPLLNSSM
                     CLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCENLTTGKHCETCISGFYGDPTNGG
                     KCQPCKCNGHASLCNTNTGKCFCTTKGVKGDECQLCEVENRYQGNPLRGTCYYTLLID
                     YQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASKNFNLNITWAASFSAGTQAG
                     EEMPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTWPIKIQVQTEQ"
ORIGIN

Alignment Scores:
Pred. No.:                0              Length:          3597
Score:                    6612.00        Matches:         1191
Percent Similarity:       93.72%         Conservative:    2
Best Local Similarity:    93.56%         Mismatches:      5
Query Match:              83.54%         Indels:          76
DB:                       9              Gaps:            1

US-09-787-097-12 (1-1429) x AF034957 (1-3597)

Qy          1 MetValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 60

Qy         21 AlaLeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArg 40
              ||||||||||||||||||||||||||||||
Db         61 GCGCTCGCGGGCAGGAGCGGCGGGCCGCAC---------------------------- 90

Qy         41 ProGlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSerProProLeuArgProArg 60

Db         90 ------------------------------------------------------------ 90

Qy         61 LeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeuProCysGlu 80

Db         90 ------------------------------------------------------------ 90

Qy         81 AlaGluAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGlu 100

Db         90 ------------------------------------------------------------ 90

Qy        101 CysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCys 120
                           ||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 ------------TGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGC 138

Qy        121 ProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGly 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 CCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGA 198

Qy        141 SerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrp 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 TCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGG 258

Qy        161 LeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 CTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAG 318

Qy        181 CysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAla 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 TGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCT 378

Qy        201 AlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGluValVal 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 GCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTT 438

Qy        221 AlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGly 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        439 GCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGA 498

Qy        241 PheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCys 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 TTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGT 558

Qy        261 LysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsnTrpLysGlyGlu 280
```

```
Db      559 AAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAA 618
Qy      281 AlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsn 300
Db      619 GCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAAT 678
Qy      301 SerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerVal 320
Db      679 TCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTT 738
Qy      321 ProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuPro 340
Db      739 CCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCC 798
Qy      341 ArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMet 360
Db      799 AGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATG 858
Qy      361 PheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeu 380
Db      859 TTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTT 918
Qy      381 ProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyr 400
Db      919 CCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATAC 978
Qy      401 LysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsnValThrAsnGlu 420
Db      979 AAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAATGAG 1038
Qy      421 LeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGlu 440
Db     1039 TTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGGAG 1098
Qy      441 GlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValVal 460
Db     1099 CAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTC 1158
Qy      461 MetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyr 480
Db     1159 ATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATAT 1218
Qy      481 AspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGly 500
Db     1219 GATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGT 1278
Qy      501 TyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyr 520
Db     1279 TACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTAC 1338
Qy      521 LysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAsp 540
Db     1339 AAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGAT 1398
Qy      541 ThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaVal 560
Db     1399 ACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTG 1458
Qy      561 IleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSer 580
Db     1459 ATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGC 1518
Qy      581 HisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrp 600
Db     1519 CATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGG 1578
Qy      601 SerValLeuProArgProAspLeuHisHisAspValAsnArgPheGlyHisSer-AlaVa 620
Db     1579 TCAGTGCTTCCCAGACCTGATTC-CACCATGATGTCAACAGATTTGGCCATTCCAGCAGT 1637
Qy      620 lLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIl 640
```

```
Db     1638 CTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACAT 1697
Qy      640 eLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAl 660
Db     1698 CCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGC 1757
Qy      660 aGlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLe 680
Db     1758 AGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCT 1817
Qy      680 uAlaThrAspGluGlnGluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAs 700
Db     1818 GGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGA 1877
Qy      700 pHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCy 720
Db     1878 CCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGACTG 1937
Qy      720 sHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSe 740
Db     1938 CCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTC 1997
Qy      740 rIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysTh 760
Db     1998 CATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAAGAC 2057
Qy      760 rSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCy 780
Db     2058 CAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTG 2117
Qy      780 sIleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLe 800
Db     2118 CATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTT 2177
Qy      800 uLysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAs 820
Db     2178 GAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAA 2237
Qy      820 nAlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuAr 840
Db     2238 TGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCG 2297
Qy      840 gIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuAr 860
Db     2298 AATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCG 2357
Qy      860 gLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLe 880
Db     2358 GAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACT 2417
Qy      880 uGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSe 900
Db     2418 ACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAG 2477
Qy      900 rThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluAr 920
Db     2478 TACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAG 2537
Qy      920 gProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGl 940
Db     2538 GCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGG 2597
Qy      940 yAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAs 960
Db     2598 AGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGA 2657
Qy      960 pSerAsnAlaTyrValAlaSerProProPheGlyGlnCysMetGluTrpTyrThrMetSe 980
Db     2658 CTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAG 2717
Qy      980 rThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPr 1000
Db     2718 CACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACC 2777
```

```
Qy       1000 oGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySe 1020
Db       2778 AGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTC 2837

Qy       1020 rTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnPr 1040
Db       2838 CTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCC 2897

Qy       1040 oLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCy 1060
Db       2898 CCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTG 2957

Qy       1060 sProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCy 1080
Db       2958 TCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTG 3017

Qy       1080 sGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspPr 1100
Db       3018 TGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCC 3077

Qy       1100 oThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnTh 1120
Db       3078 CACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACAC 3137

Qy       1120 rAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCy 1140
Db       3138 CAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATG 3197

Qy       1140 sGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLe 1160
Db       3198 TGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCT 3257

Qy       1160 uIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIl 1180
Db       3258 TATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTAT 3317

Qy       1180 eAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSe 1200
Db       3318 CAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTC 3377

Qy       1200 rLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGl 1220
Db       3378 CAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGG 3437

Qy       1220 yGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAs 1240
Db       3438 AGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAA 3497

Qy       1240 nGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPh 1260
Db       3498 TGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTT 3557

Qy       1260 eThrTrpProIleLysIleGlnIleAlaPheSerGln 1272
Db       3558 CACCTGGCCCATCAAAATTCAGGTGCAAACTGAACAA 3594
```

```
RESULT 11
AB038388
LOCUS       AB038388               4500 bp    mRNA    linear   ROD 18-JAN-2001
DEFINITION  Rattus norvegicus Atrn mRNA for attractin, complete cds.
ACCESSION   AB038388
VERSION     AB038388.1  GI:12275311
KEYWORDS    attractin.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (sites)
  AUTHORS   Kuramoto,T., Kitada,K., Inui,T., Sasaki,Y., Ito,K., Hase,T.,
            Kawaguchi,S., Ogawa,Y., Nakao,K., Barsh,G.S., Nagao,M., Ushijima,T.
            and Serikawa,T.
  TITLE     Attractin/mahogany/zitter plays a critical role in myelination of
            the central nervous system
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (2), 559-564 (2001)
  MEDLINE   21143347
   PUBMED   11209055
REFERENCE   2  (bases 1 to 4500)
  AUTHORS   Kuramoto,T., Serikawa,T. and Ushijima,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-FEB-2000) Takashi Kuramoto, National Cancer Center
            Research Institute, Carcinogenesis Division; Tsukiji 5-1-1,
            Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramot@gan2.ncc.go.jp,
            Tel:81-33542-2511(ex.4521), Fax:81-35565-1753)
FEATURES             Location/Qualifiers
     source          1. .4500
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="WTC"
                     /db_xref="taxon:10116"
     gene            1. .4500
                     /gene="Atrn"
     CDS             61. .3888
                     /gene="Atrn"
                     /codon_start=1
                     /product="attractin"
                     /protein_id="BAB21018.1"
                     /db_xref="GI:12275312"
                     /translation="MVAAAAAAAEATEARLRGYTTATAAPAGWKERQHRPCAATGAWRP
                     WPRAGLCLPRVLSRALSPPPLLPLLPLLFSLLLLPLPREAEAAAVAAAVSGSAAAEAK
                     ECDRPCVNGGRCNPGTGQCVCPTGWVGEQCQHCGGRFRLTGSSGFVTDGPGNYKYKTK
                     CTWLIEGQPNKIMRLRFNHFATECSWDHLYVYDGDSIYAPLIAAFSGLIVPERDGNET
                     APEVTVTSGYALLHFFSDAAYNLTGFNITYNFDMCPNNCSGRGECKSSNSSSTVECEC
                     SENWKGESCDIPHCTDNCGFPHRGICNASDTRGCSCFPHWQGPGCSIPVPANQSFWTR
                     EEYSDLKLPRASHKAEVNGNIMWVVGGYMFNHSDYSMVLAYDLASREWLSLNHSVNSV
                     VVRYGHSLALHKDKIYMYGGKIDSTGNVTNELRVFHIHNESWVLLTPKAKDQYAVVGH
                     SAHIVTLSSGRVVMLVIFGHCPLYGYISVVQEYDLEKNTWSILQTQGALVQGGYGHSS
                     VYDHRTKALYVHGGYKAFSANKYRLADDLYRYHVDTQMWTILKDSRFFRYLHTAVIVS
                     GTMLVFGGNTHNDTSMSHGAKCFSSDFMAYDIACDRWSVLPRPELHHDVNRFGHSAVL
                     HNSTMYVFGGFNSLLLSDVLVFTSEQCDAHRSEAACVAAGPGIRCLWDTQSSRCTSWE
                     LATEEQAEKLKSECFSKRTLDHDRCDQHTDCYSCTANTNDCHWCNDHCVPVNHSCTEG
                     QISIAKYDNCPKDNPMYYCNKKTSCRSCALDQNCQWEPRNQECIALPENICGIGWHLV
                     GNSCLKITTAKENYDNAKLSCRNHNAFLASLTSQKKVEFVLKQLRLMQSSQSTSKLTL
                     TPWVGLRKINVSYWCWEDMSPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCIN
                     PLNGSVCERPANHSAKQCRTPCALRTACGECTSSSSECMWCSNMKQCVDSNAYVASFP
                     FGQCMEWYTMSSCPPENCSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKM
                     PSHASTGNVYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCEDLTT
                     GKHCETCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGEECQLCEVE
                     NRYQGNPLKGTCYYTLLIDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASK
                     NFNLNITWATSFPAGTQTGEEVPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSN
                     FTWPIKIQVRVTS"
ORIGIN
```

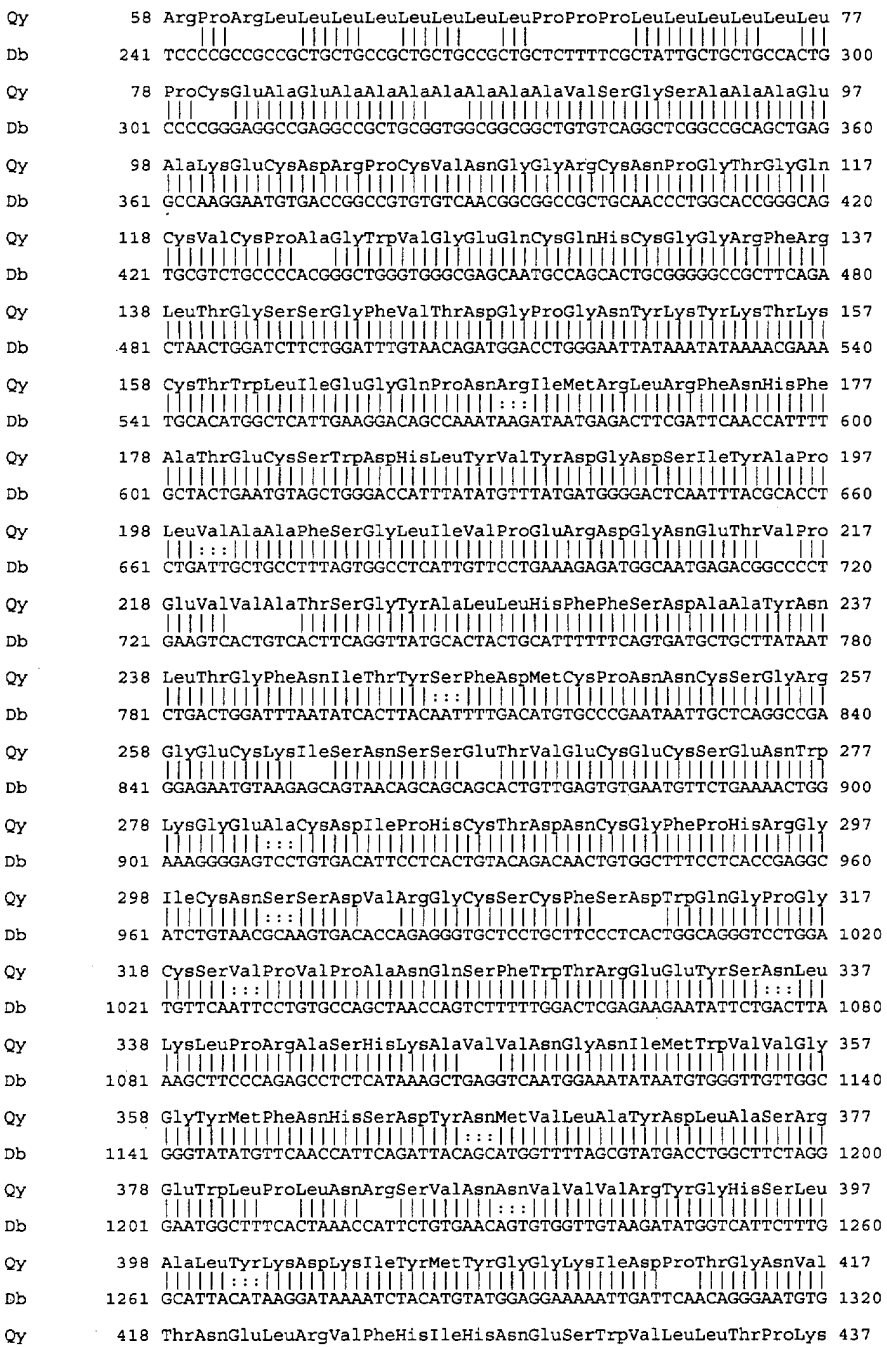
```
Alignment Scores:
Pred. No.:                0                Length:          4500
Score:                    6509.50          Matches:         1170
Percent Similarity:       94.29%           Conservative:    36
Best Local Similarity:    91.48%           Mismatches:      68
Query Match:              82.24%           Indels:          5
DB:                       10               Gaps:            2

US-09-787-097-12 (1-1429) x AB038388 (1-4500)

Qy          1 MetValAlaAlaAlaAla---------AlaThrGluAlaArgLeuArgArgArgThrAla 17
Db         61 ATGGTGGCGGCGGCGGCGGCGGCGGAGGCGACTGAGGCGCGGTTGAGGGGGTACACGACG 120

Qy         18 AlaThrAlaAlaLeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArg 37
Db        121 GCGACAGCAGCGCCCGCGGGCTGGAAGGAGAGGCAGCACCGACCCTGCGCCGCGACAGGG 180

Qy         38 AlaGlyArgProGlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSerProProLeu 57
Db        181 GCCTGGAGGCCGTGGCCGCGCGCCGGGCTGTGTCTCCCGCGGGTGCTGTCGCGGGCGCTA 240
```


Qy 58 ArgProArgLeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeuLeu 77
Db 241 TCCCCGCCGCCGCTGCTGCCGCTGCTGCCGCTGCTCTTTTCGCTATTGCTGCTGCCACTG 300
Qy 78 ProCysGluAlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAlaGlu 97
Db 301 CCCCGGGAGGCCGAGGCCGCTGCGGTGGCGGCGGCTGTGTCAGGCTCGGCCGCAGCTGAG 360
Qy 98 AlaLysGluCysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGlyGln 117
Db 361 GCCAAGGAATGTGACCGGCCGTGTGTCAACGGCGGCCGCTGCAACCCTGGCACCGGGCAG 420
Qy 118 CysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPheArg 137
Db 421 TGCGTCTGCCCCACGGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGA 480
Qy 138 LeuThrGlySerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLys 157
Db 481 CTAACTGGATCTTCTGGATTTGTAACAGATGGACCTGGGAATTATAAATATAAAACGAAA 540
Qy 158 CysThrTrpLeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPhe 177
Db 541 TGCACATGGCTCATTGAAGGACAGCCAAATAAGATAATGAGACTTCGATTCAACCATTTT 600
Qy 178 AlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaPro 197
Db 601 GCTACTGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATTTACGCACCT 660
Qy 198 LeuValAlaAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValPro 217
Db 661 CTGATTGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCCCCT 720
Qy 218 GluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsn 237
Db 721 GAAGTCACTGTCACTTCAGGTTATGCACTACTGCATTTTTTCAGTGATGCTGCTTATAAT 780
Qy 238 LeuThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArg 257
Db 781 CTGACTGGATTTAATATCACTTACAATTTTGACATGTGCCCGAATAATTGCTCAGGCCGA 840
Qy 258 GlyGluCysLysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsnTrp 277
Db 841 GGAGAATGTAAGAGCAGTAACAGCAGCAGCACTGTTGAGTGTGAATGTTCTGAAAACTGG 900
Qy 278 LysGlyGluAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGly 297
Db 901 AAAGGGGAGTCCTGTGACATTCCTCACTGTACAGACAACTGTGGCTTTCCTCACCGAGGC 960
Qy 298 IleCysAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyProGly 317
Db 961 ATCTGTAACGCAAGTGACACCAGAGGGTGCTCCTGCTTCCCTCACTGGCAGGGTCCTGGA 1020
Qy 318 CysSerValProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeu 337
Db 1021 TGTTCAATTCCTGTGCCAGCTAACCAGTCTTTTTGGACTCGAGAAGAATATTCTGACTTA 1080
Qy 338 LysLeuProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValValGly 357
Db 1081 AAGCTTCCCAGAGCCTCTCATAAAGCTGAGGTCAATGGAAATATAATGTGGGTTGTTGGC 1140
Qy 358 GlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArg 377
Db 1141 GGGTATATGTTCAACCATTCAGATTACAGCATGGTTTTAGCGTATGACCTGGCTTCTAGG 1200
Qy 378 GluTrpLeuProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSerLeu 397
Db 1201 GAATGGCTTTCACTAAACCATTCTGTGAACAGTGTGGTTGTAAGATATGGTCATTCTTTG 1260
Qy 398 AlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsnVal 417
Db 1261 GCATTACATAAGGATAAAATCTACATGTATGGAGGAAAAATTGATTCAACAGGGAATGTG 1320
Qy 418 ThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrProLys 437
Db 1321 ACCAACGAGCTGAGAGTATTTCATATTCATAATGAATCATGGGTATTGTTAACCCCCAAA 1380
Qy 438 AlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsnGly 457
Db 1381 GCTAAGGATCAGTATGCAGTGGTTGGACACTCAGCACACATTGTTACACTGTCATCTGGC 1440
Qy 458 ArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsnVal 477
Db 1441 CGTGTGGTCATGTTGGTCATCTTCGGTCATTGTCCGCTCTACGGATATATAAGTGTTGTG 1500
Qy 478 GlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuVal 497
Db 1501 CAGGAATATGACTTGGAAAAGAACACATGGAGTATATTACAAACTCAAGGTGCTCTAGTG 1560
Qy 498 GlnGlyGlyTyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrValHis 517
Db 1561 CAAGGGGGTTATGGCCACAGCAGTGTTTATGATCACAGGACCAAGGCTCTGTATGTTCAT 1620
Qy 518 GlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyr 537
Db 1621 GGTGGCTACAAGGCTTTCAGCGCCAACAAGTACCGGCTTGCAGATGACCTCTACAGATAC 1680
Qy 538 AspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeuHis 557
Db 1681 CATGTGGACACTCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCAT 1740
Qy 558 ThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThr 577
Db 1741 ACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTCGGAGGGAACACACACAATGACACT 1800
Qy 578 SerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAlaCys 597
Db 1801 TCCATGAGCCACGGTGCCAAATGCTTCTCCTCAGACTTCATGGCTTATGACATTGCTTGT 1860
Qy 598 AspArgTrpSerValLeuProArgProAspLeuHisHisAspValAsnArgPheGlyHis 617
Db 1861 GACCGATGGTCAGTTCTCCCCAGACCTGAGCTCCATCACGATGTCAACCGATTTGGCCAC 1920
Qy 618 SerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeu 637
Db 1921 TCAGCAGTCTTGCACAACAGCACCATGTATGTGTTCGGCGGCTTCAACAGCCTCCTCCTC 1980
Qy 638 SerAspIleLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCys 657
Db 1981 AGTGACGTCTTAGTCTTCACCTCGGAGCAGTGTGATGCACACCGCAGCGAAGCTGCTTGT 2040
Qy 658 LeuAlaAlaGlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIleSer 677
Db 2041 GTGGCAGCAGGACCTGGCATCCGGTGTCTGTGGGACACACAGTCGTCTCGATGTACCTCC 2100
Qy 678 TrpAlaLeuAlaThrAspGluGlnGluGluLysLeuLysSerGluCysPheSerLysArg 697
Db 2101 TGGGAGTTGGCAACTGAAGAACAAGCAGAAAAGTTAAAATCAGAATGTTTCTCTAAAAGA 2160
Qy 698 ThrLeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThr 717
Db 2161 ACCCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAATACC 2220
Qy 718 AsnAspCysHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGluGly 737
Db 2221 AATGACTGCCACTGGTGCAATGATCACTGTGTCCCTGTGAACCACAGCTGCACAGAAGGC 2280
Qy 738 GlnIleSerIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsn 757
Db 2281 CAGATCTCCATTGCCAAGTATGACAATTGCCCCAAGGATAACCCCATGTACTACTGCAAT 2340
Qy 758 LysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsn 777
Db 2341 AAGAAAACCAGCTGCAGGAGCTGTGCCCTAGATCAGAATTGCCAGTGGGAACCTCGAAAT 2400
Qy 778 GlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsn 797

```
Db       2401 CAAGAGTGCATCGCCCTGCCCGAAAACATCTGTGGCATTGGCTGGCATTTGGTCGGAAAC 2460

Qy        798 SerCysLeuLysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArg 817

Db       2461 TCGTGTCTGAAAATCACTACTGCTAAGGAAAATTATGACAATGCTAAATTGTCCTGTAGA 2520

Qy        818 AsnHisAsnAlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLys 837

Db       2521 AACCACAATGCCTTTTTGGCTTCGCTTACGTCCCAGAAGAAGGTGGAGTTTGTCCTTAAG 2580

Qy        838 GlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpVal 857

Db       2581 CAGCTTCGATTAATGCAATCATCTCAGAGCACGTCCAAGCTCACTCTGACTCCATGGGTT 2640

Qy        858 GlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsn 877

Db       2641 GGTCTTCGGAAGATCAATGTATCTTACTGGTGCTGGGAGGATATGTCTCCATTCACAAAT 2700

Qy        878 SerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSer 897

Db       2701 AGTTTGCTGCAGTGGATGCCATCTGAGCCCAGTGATGCCGGCTTCTGTGGGATCTTGTCA 2760

Qy        898 GluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerVal 917

Db       2761 GAGCCTAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGAAGTGTT 2820

Qy        918 CysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThr 937

Db       2821 TGTGAAAGGCCTGCAAACCACAGTGCCAAGCAGTGCCGGACACCCTGTGCCCTGCGGACA 2880

Qy        938 AlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGln 957

Db       2881 GCATGTGGTGAGTGCACCAGCAGCAGCTCCGAATGCATGTGGTGCAGCAACATGAAGCAG 2940

Qy        958 CysValAspSerAsnAlaTyrValAlaSerProProPheGlyGlnCysMetGluTrpTyr 977

Db       2941 TGCGTAGACTCCAATGCCTACGTAGCCTCCTTTCCTTTCGGCCAGTGTATGGAGTGGTAT 3000

Qy        978 ThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeu 997

Db       3001 ACAATGAGCAGCTGCCCACCTGAAAATTGCTCTGGCTACTGTACCTGCAGCCATTGCTTG 3060

Qy        998 GluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIle 1017

Db       3061 GAGCAGCCAGGCTGTGGTTGGTGTACTGATCCTAGCAATACTGGGAAAGGCAAATGTATT 3120

Qy       1018 GluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyr 1037

Db       3121 GAGGGAAGCTATAAAGGACCCGTGAAGATGCCTTCACATGCCTCTACAGGAAATGTGTAC 3180

Qy       1038 ProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPhe 1057

Db       3181 CCACAGCCCCTTCTGAACTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTT 3240

Qy       1058 IleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCys 1077

Db       3241 ATTCACTGTCCAGCTTGCCAGTGCAATGGACACAGCAAGTGCATCAACCAGAGCATCTGT 3300

Qy       1078 GluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyr 1097

Db       3301 GAGAAGTGTGAGGACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTAT 3360

Qy       1098 GlyAspProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeu 1117

Db       3361 GGTGACCCGACTAATGGAGGCAAATGTCAGCCATGCAAGTGCAACGGGCACGCATCACTG 3420

Qy       1118 CysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCys 1137

Db       3421 TGCAACACCAACACTGGCAAGTGCTTCTGTACCACCAAGGGCGTCAAGGGAGAGGAGTGC 3480

Qy       1138 GlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyr 1157

Db       3481 CAGCTATGTGAGGTAGAAAATCGATACCAGGGAAACCCTCTCAAAGGAACATGTTACTAT 3540

Qy       1158 ThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyr 1177

Db       3541 ACCCTTCTCATTGACTATCAGTTCACTTTTAGCCTGTCCCAGGAAGACGACCGCTATTAC 3600

Qy       1178 ThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIle 1197

Db       3601 ACAGCCATCAATTTTGTGGCTACTCCTGATGAACAAAACAGGGATTTGGACATGTTCATC 3660

Qy       1198 AsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThr 1217

Db       3661 AATGCCTCCAAAAACTTCAACCTCAACATCACCTGGGCCACCAGTTTCCCAGCTGGAACC 3720

Qy       1218 GlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSer 1237

Db       3721 CAGACTGGAGAAGAGGTGCCTGTTGTTTCAAAAACCAACATCAAGGAATACAAAGATAGC 3780

Qy       1238 PheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrVal 1257

Db       3781 TTCTCTAATGAGAAATTTGATTTTCGCAACCATCCAAACATTACTTTCTTTGTTTATGTC 3840

Qy       1258 SerAsnPheThrTrpProIleLysIleGlnIleAlaPheSer------GlnHisSer 1274

Db       3841 AGTAATTTCACTTGGCCCATCAAAATTCAAGTAAGAGTTACTTCTTAACGTCATTCG 3897
```

```
RESULT 12
AB093245
LOCUS       AB093245                5683 bp    mRNA    linear   ROD 07-FEB-2003
DEFINITION  Mus musculus mRNA for mKIAA0548 protein.
ACCESSION   AB093245
VERSION     AB093245.1  GI:26006172
KEYWORDS    FLI_CDNA.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Okazaki,N., Kikuno,R., Ohara,R., Inamoto,S., Hara,Y., Nagase,T.,
            Ohara,O. and Koga,H.
  TITLE     Prediction of the coding sequences of mouse homologues of KIAA
            gene: I. The complete nucleotide sequences of 100 mouse
            KIAA-homologous cDNAs identified by screening of terminal sequences
            of cDNA clones randomly sampled from size-fractionated libraries
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 5683)
  AUTHORS   Okazaki,N., Kikuno,R., Nagase,T., Ohara,O. and Koga,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-OCT-2002) Hisashi Koga, Kazusa DNA Research
            Institute, Laboratory for Genome Informatics; 2-6-7
            Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan
            (E-mail:mouse@kazusa.or.jp, Tel:81-438-52-3919, Fax:81-438-52-3918)
COMMENT     The CREATE program supported by Japan science and technology
            corporation; cDNA full insert sequencing: Kazusa DNA Research
            Institute; cDNA library construction, clone selection and 5'- &
            3'-end one pass sequencing.
FEATURES             Location/Qualifiers
     source          1. .5683
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="mbg03041"
                     /tissue_type="brain"
                     /dev_stage="adult"
                     /note="vector:modified pBC SK+"
     gene            1. .5683
                     /gene="mKIAA0548"
     CDS             <1. .3181
                     /gene="mKIAA0548"
                     /note="CDS is predicted by in silico analysis. Start codon
                     is not identified."
                     /codon_start=2
                     /evidence=not_experimental
                     /product="mKIAA0548 protein"
```

```
                     /protein_id="BAC41429.1"
                     /db_xref="GI:26006173"
                     /translation="AYDLTSREWLPLNHSVNSVVVRYGHSLALHKDKIYMYGGKIDST
                     GNVTNELRVFHIHNESWVLLTPKAKDQYAVVGHSAHIVTLASGRVVMLVIFGHCPLYG
                     YISVVQEYDLEKNTWSILHTQGALVQGGYGHSSVYDDRTKALYVHGGYKAFSANKYRL
                     ADDLYRYDVDTQMWTILKDSRFFRYLHTAVIVSGTMLVFGGNTHNDTSMSHGAKCFSS
                     DFMAYDIACDRWSVLPRPELHHDVNRFGHSAVLYNSTMYVFGGFNSLLLSDVLVFTLE
                     QCDAHRSEAACVAAGPGIRCLWDTQSSRCTSWELATEEQAEKLKSECFSKRTLDHDRC
                     DQHTDCYSCTANTNDCHWCNDHCVPVNHSCTEGQISIAKYESCPKDNPMYYCNKKTSC
                     RSCALDQNCQWEPRNQECIALPENICGNGWHLVGNSCLKITTAKENYDNAKLSCRNHN
                     AFLASLTSQKKVEFVLKQLRLMQSSQSMSKLTLTPWVGLRKINVSYWCWEDMSPFTNS
                     LLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLNGSVCERPANHSAKQCRTPCALR
                     TACGECTSSSSECMWCSNMKQCVDSNAYVASFPFGQCMEWYTMSSCPPENCSGYCTCS
                     HCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKMPSQASAGNVYPQPLLNSSMCLEDSR
                     YNWSFIHCPACQCNGHSKCINQSVCEKCEDLTTGKHCETCISGFYGDPTNGGKCQPCK
                     CNGHASLCNTNTGKCFCTTKGVKGDECQLCEVENRYQGNPLKGTCYYTLLIDYQFTFS
                     LSQEDDRYYTAINFVATPDEQNRDLDMFINASKNFNLNITWATSFPAGTQTGEEVPVV
                     SKTTIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTWPIKIQIAFSQHSNFMDLVQFFV
                     TFFSCFLSLLLVAAVVWKIKQSCWASRRREQLLREMQQMASRPFASVNVALETDEEPP
                     DLIGGSIKTVPKPIALEPCFGNKAAVLSVFVRLPRGLGGIPPPGQSGLAVASALVDIS
                     QQMPIVYKEKSGAVRNRKQQPPAQPGTCI"
ORIGIN

Alignment Scores:
Pred. No.:                0              Length:          5683
Score:                    5624.00        Matches:         1010
Percent Similarity:       97.73%         Conservative:    25
Best Local Similarity:    95.37%         Mismatches:      24
Query Match:              71.05%         Indels:          0
DB:                       10             Gaps:            0

US-09-787-097-12 (1-1429) x AB093245 (1-5683)

Qy        371 AlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSerValAsnAsnValVal 390
Db          2 GCGTATGACCTGACTTCTAGGGAATGGCTTCCACTAAACCATTCTGTGAACAGTGTGGTT 61

Qy        391 ValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLys 410
Db         62 GTAAGATATGGTCATTCTTTGGCATTACATAAGGATAAAATCTACATGTATGGAGGAAAA 121

Qy        411 IleAspProThrGlyAsnValThrAsnGluLeuArgValPheHisIleHisAsnGluSer 430
Db        122 ATTGATTCAACAGGGAACGTGACCAATGAGCTGAGAGTATTTCATATTCATAATGAATCA 181

Qy        431 TrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHis 450
Db        182 TGGGTATTGTTAACTCCGAAAGCTAAGGATCAGTATGCAGTGGTTGGACACTCAGCACAC 241

Qy        451 IleValThrLeuLysAsnGlyArgValValMetLeuValIlePheGlyHisCysProLeu 470
Db        242 ATTGTTACACTGGCATCTGGCCGTGTGGTCATGTTGGTCATCTTCGGTCATTGCCCACTC 301

Qy        471 TyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeu 490
Db        302 TATGGATATATAAGCGTTGTGCAGGAATATGACTTGGAAAAGAACACATGGAGTATATTA 361

Qy        491 HisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisSerSerValTyrAspHisArg 510
Db        362 CATACTCAGGGTGCTCTTGTGCAAGGGGGTTATGGCCACAGTAGTGTTTATGATGACAGG 421

Qy        511 ThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeu 530
Db        422 ACCAAGGCTCTGTACGTTCATGGTGGCTACAAGGCTTTCAGCGCCAACAAATACCGGCTT 481

Qy        531 AlaAspAspLeuTyrArgTyrAspValAspThrGlnMetTrpThrIleLeuLysAspSer 550
Db        482 GCAGATGACCTCTACAGATACGATGTGGATACTCAGATGTGGACCATTCTTAAGGACAGC 541

Qy        551 ArgPhePheArgTyrLeuHisThrAlaValIleValSerGlyThrMetLeuValPheGly 570
Db        542 CGATTTTTCCGTTACTTGCATACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGA 601

Qy        571 GlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysPheSerSerAspPhe 590
Db        602 GGGAACACACACAATGACACTTCCATGAGCCACGGTGCCAAATGCTTCTCCTCGGACTTC 661

Qy        591 MetAlaTyrAspIleAlaCysAspArgTrpSerValLeuProArgProAspLeuHisHis 610
Db        662 ATGGCTTATGACATTGCTTGTGACCGATGGTCAGTGCTTCCCAGACCTGAGCTCCATCAT 721

Qy        611 AspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrMetTyrValPheGly 630
Db        722 GATGTCAACAGATTTGGCCATTCAGCAGTCTTGTACAACAGCACCATGTATGTGTTCGGC 781

Qy        631 GlyPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSerGluGlnCysAspAla 650
Db        782 GGCTTCAACAGCCTCCTCCTCAGTGACGTCTTGGTCTTTACCTTGGAGCAGTGCGATGCA 841

Qy        651 HisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgCysValTrpAsnThr 670
Db        842 CACCGCAGTGAAGCTGCTTGTGTGGCAGCAGGACCTGGTATCCGGTGTCTGTGGGACACA 901

Qy        671 GlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnGluGluLysLeuLys 690
Db        902 CAGTCGTCTCGATGTACCTCCTGGGAGTTGGCAACTGAAGAACAAGCAGAAAAGTTAAAA 961

Qy        691 SerGluCysPheSerLysArgThrLeuAspHisAspArgCysAspGlnHisThrAspCys 710
Db        962 TCAGAGTGTTTTTCTAAAAGAACCCTTGACCATGACAGATGTGACCAGCACACAGATTGT 1021

Qy        711 TyrSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAspHisCysValProArg 730
Db       1022 TACAGCTGCACAGCCAATACCAATGACTGCCACTGGTGCAATGATCACTGTGTCCCTGTG 1081

Qy        731 AsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGluAsnCysProLysAsp 750
Db       1082 AACCACAGCTGCACAGAAGGCCAGATCTCCATTGCCAAGTATGAGAGTTGCCCCAAGGAT 1141

Qy        751 AsnProMetTyrTyrCysAsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsn 770
Db       1142 AACCCCATGTACTACTGCAATAAGAAAACCAGCTGCAGGAGCTGTGCCCTAGACCAGAAC 1201

Qy        771 CysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuProGluAsnIleCysGlyIle 790
Db       1202 TGCCAGTGGGAGCCCCGGAATCAAGAGTGCATCGCCCTGCCGGAAAATATCTGTGGCAAT 1261

Qy        791 GlyTrpHisLeuValGlyAsnSerCysLeuLysIleThrThrAlaLysGluAsnTyrAsp 810
Db       1262 GGCTGGCATTTGGTTGGAAACTCGTGTCTGAAAATCACTACTGCTAAGGAGAATTATGAC 1321

Qy        811 AsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSerLeuThrThrGlnLys 830
Db       1322 AATGCTAAATTGTCCTGTAGGAACCACAATGCCTTTTTGGCTTCCCTCACATCCCAGAAG 1381

Qy        831 LysValGluPheValLeuLysGlnLeuArgIleMetGlnSerSerGlnSerMetSerLys 850
Db       1382 AAGGTGGAGTTTGTCCTTAAGCAGCTTCGATTAATGCAATCATCTCAAAGTATGTCCAAG 1441

Qy        851 LeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCysTrpGlu 870
Db       1442 CTCACTCTGACTCCATGGGTTGGTCTTCGGAAGATCAATGTGTCTTACTGGTGCTGGGAG 1501

Qy        871 AspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSerAspAla 890
Db       1502 GATATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGTGATGCT 1561

Qy        891 GlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThrCysIle 910
Db       1562 GGCTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGATTAAAGGCTGCAACCTGCATC 1621

Qy        911 AsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSerAlaLysGlnCysArg 930
Db       1622 AACCCTCTCAATGGCAGCGTCTGTGAAAGGCCTGCAAACCACAGTGCCAAGCAGTGCCGG 1681

Qy        931 ThrProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGlySerSerGluCysMet 950
```

```
Db      1682 ACACCATGTGCCCTGCGGACAGCGTGTGGCGAGTGCACTAGCAGCAGCTCGGAGTGCATG 1741
Qy       951 TrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrValAlaSerProPhePhe 970
Db      1742 TGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTACGTGGCCTCCTTCCCTTTT 1801
Qy       971 GlyGlnCysMetGluTrpTyrThrMetSerThrCysProProGluAsnCysSerGlyTyr 990
Db      1802 GGCCAGTGTATGGAATGGTATACGATGAGCAGCTGCCCACCTGAAAATTGCTCTGGCTAT 1861
Qy       991 CysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCysThrAspProSerAsn 1010
Db      1862 TGTACCTGCAGCCATTGCTTGGAGCAGCCAGGCTGTGGTTGGTGTACTGATCCTAGCAAT 1921
Qy      1011 ThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyProValLysMetProSerGln 1030
Db      1922 ACTGGGAAAGGAAAATGTATTGAGGGCAGCTATAAAGGACCTGTGAAGATGCCGTCACAG 1981
Qy      1031 AlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerSerMetCysLeuGluAsp 1050
Db      1982 GCCTCTGCAGGAAATGTGTATCCACAGCCCCTTCTGAACTCCAGCATGTGTCTAGAGGAC 2041
Qy      1051 SerArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnCysAsnGlyHisSerLys 1070
Db      2042 AGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAGTGCAACGGACACAGCAAA 2101
Qy      1071 CysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrThrGlyLysHisCysGlu 1090
Db      2102 TGCATCAACCAGAGTGTCTGTGAGAAGTGTGAGGACCTGACCACGGGCAAGCACTGCGAG 2161
Qy      1091 ThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLys 1110
Db      2162 ACCTGCATATCTGGCTTCTATGGTGACCCGACTAATGGAGGCAAATGTCAGCCATGCAAG 2221
Qy      1111 CysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLys 1130
Db      2222 TGCAATGGGCACGCATCACTGTGCAACACCAACACCGGCAAGTGCTTCTGTACCACCAAA 2281
Qy      1131 GlyValLysGlyAspGluCysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnPro 1150
Db      2282 GGTGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAGGGAAATCCT 2341
Qy      1151 LeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSer 1170
Db      2342 CTCAAAGGAACATGCTACTATACCCTTCTCATTGACTATCAGTTCACCTTTAGCCTGTCC 2401
Qy      1171 GlnGluAspAspArgTyrTyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsn 1190
Db      2402 CAGGAAGACGACCGCTACTACACAGCCATCAACTTTGTGGCTACTCCTGATGAACAAAAC 2461
Qy      1191 ArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAla 1210
Db      2462 AGGGATTTGGACATGTTCATCAATGCCTCCAAAAACTTCAACCTCAACATCACCTGGGCC 2521
Qy      1211 AlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProValValSerLysThrAsn 1230
Db      2522 ACCAGCTTCCCAGCCGGAACCCAGACTGGAGAAGAGGTGCCTGTTGTTTCAAAAACCACC 2581
Qy      1231 IleLysGluTyrLysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsn 1250
Db      2582 ATCAAGGAATACAAAGATAGCTTCTCTAATGAGAAATTTGATTTTCGCAACCATCCAAAC 2641
Qy      1251 IleThrPhePheValTyrValSerAsnPheThrTrpProIleLysIleGlnIleAlaPhe 1270
Db      2642 ATCACTTTCTTTGTTTATGTCAGTAATTTCACTTGGCCCATCAAAATTCAGATTGCCTTC 2701
Qy      1271 SerGlnHisSerAsnPheMetAspLeuValGlnPhePheValThrPhePheSerCysPhe 1290
Db      2702 TCCCAGCACAGCAACTTCATGGACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTC 2761
Qy      1291 LeuSerLeuLeuLeuValAlaAlaValValTrpLysIleLysGlnSerCysTrpAlaSer 1310
```

```
Db      2762 CTCTCGCTGCTTCTGGTGGCTGCAGTGGTCTGGAAGATCAAGCAGAGCTGTTGGGCATCC 2821
Qy      1311 ArgArgArgGluGlnLeuLeuArgGluMetGlnGlnMetAlaSerArgProPheAlaSer 1330
Db      2822 AGGCGGAGAGAGCAACTTCTTCGGGAGATGCAACAGATGGCCAGCCGCCCCTTTGCTTCT 2881
Qy      1331 ValAsnValAlaLeuGluThrAspGluGluProProAspLeuIleGlyGlySerIleLys 1350
Db      2882 GTAAACGTTGCCTTGGAAACAGATGAAGAACCTCCTGATCTCATTGGGGGAAGTATAAAG 2941
Qy      1351 ThrValProLysProIleAlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSer 1370
Db      2942 ACCGTTCCTAAGCCCATTGCCCTGGAGCCCTGCTTTGGTAACAAAGCCGCAGTCCTCTCT 3001
Qy      1371 ValPheValArgLeuProArgGlyLeuGlyGlyIleProProProGlyGlnSerGlyLeu 1390
Db      3002 GTATTCGTGAGGCTCCCTCGAGGACTGGGAGGAATCCCTCCTCCTGGTCAGTCAGGTCTC 3061
Qy      1391 AlaValAlaSerAlaLeuValAspIleSerGlnGlnMetProIleValTyrLysGluLys 1410
Db      3062 GCTGTGGCCAGTGCCCTGGTGGACATTTCTCAGCAGATGCCAATAGTGTACAAGGAGAAG 3121
Qy      1411 SerGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnProGlyThrCysIle 1429
Db      3122 TCAGGAGCTGTAAGAAACCGGAAGCAGCAGCCGCCTGCACAGCCTGGAACCTGCATT 3178
```

```
RESULT 13
AY442317
LOCUS       AY442317                4140 bp    mRNA    linear   PRI 15-NOV-2003
DEFINITION  Homo sapiens attractin-like protein mRNA, complete cds.
ACCESSION   AY442317
VERSION     AY442317.1  GI:38230554
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4140)
  AUTHORS   Milton,N.G.N.
  TITLE     Attractin-like peptides
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 4140)
  AUTHORS   Milton,N.G.N.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-OCT-2003) R&D, NeuroDelta Ltd., 35 Hillsborough Road,
            London SE22 8QE, UK
FEATURES             Location/Qualifiers
     source          1. .4140
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="10"
     CDS             1. .4140
                     /note="membrane receptor; similar to KIAA0534 protein"
                     /codon_start=1
                     /product="attractin-like protein"
                     /protein_id="AAR14297.1"
                     /db_xref="GI:38230555"
                     /translation="METGGRARTGTPQPAAPGVWRARPAGGGGGGASSWLLDGNSWLL
                     CYGFLYLALYAQVSQSKPCERTGSCFSGRCVNSTCLCDPGWVGDQCQHCQGRFKLTEP
                     SGYLTDGPINYKYKTKCTWLIEGYPNAVLRLRFNHFATECSWDHMYVYDGDSIYAPLI
                     AVLSGLIVPEIRGNETVPEVVTTSGYALLHFFSDAAYNLTGFNIFYSINSCPNNCSGH
                     GKCTTSVSVPSQVYCECDKYWKGEACDIPYCKANCGSPDHGYCDLTGEKLCVCNDSWQ
                     GPDCSLNVPSTESYWILPNVKPFSPSVGRASHKAVLHGKFMWVIGGYTFNYSSFQMVL
                     NYNLESSIWNVGTPSRGPLQRYGHSLALYQENIFMYGGRIETNDGNVTDELWVFNIHS
                     QSWSTKTPTVLGHGQQYAVEGHSAHIMELDSRDVVMIIIFGYSAIYGYTSSIQEYHIS
                     SNTWLVPETKGAIVQGGYGHTSVYDEITKSIYVHGGYKALPGNKYGLVDDLYKYEVNT
                     KTWTILKESGFARYLHSAVLINGAMLIFGGNTHNDTSLSNGAKCFSADFLAYDIACDE
                     WKILPKPNLHRDVNRFGHSAVVINGSMYIFGGFSSVLLNDILVYKPPNCKAFRDEELC
                     KNAGPGIKCVWNKNHCESWESGNTNNILRAKCPPKTAASDDRCYRYADCASCTANTNG
                     CQWCDDKKCISANSNCSMSVKNYTKCHVRNEQICNKLTSCKSCSLNLNCQWDQRQQEC
                     QALPAHLCGEGWSHIGDACLRVNSSRENYDNAKLYCYNLSGNLASLTTSKEVEFVLDE
                     IQKYTQQKVSPWVGLRKINISYWGWEDMSPFTNTTLQWLPGEPNDSGFCAYLERAAVA
```

GLKANPCTSMANGLVCEKPVVSPNQNARPCKKPCSLRTSCSNCTSNGMECMWCSSTKR
CVDSNAYIISFPYGQCLEWQTATCSPQNCSGLRTCGQCLEQPGCGWCNDPSNTGRGHC
IEGSSRGPMKLIGMHHSEMVLDTNLCPKEKNYEWSFIQCPACQCNGHSTCINNNVCEQ
CKNLTTGKQCQDCMPGYYGDPTNGGQCTACTCSGHANICHLHTGKCFCTTKGIKGDQC
QLCDSENRYVGNPLRGTCYYSLLIDYQFTFSLLQEDDRHHTAINFIANPEQSNKNLDI
SINASNNFNLNITWSVGSTAGTISGEETSIVSKNNIKEYRDSFSYEKFNFRSNPNITF
YVYVSNFSWPIKIQIAFSQHNTIMDLVQFFVTFFSCFLSLLLVAAVVWKIKQTCWASR
RREQLLRERQQMASRPFASVDVALEVGAEQTEFLRGPLEGAPKPIAIEPCAGNRAAVL
TVFLCLPRGSSGAPPPGQSGLAIASALIDISQQKASDSKDKTSGVRNRKHLSTRQGTC
V"

ORIGIN

Alignment Scores:

Pred. No.:	3.78e-281	Length:	4140
Score:	4638.50	Matches:	843
Percent Similarity:	73.88%	Conservative:	209
Best Local Similarity:	59.20%	Mismatches:	300
Query Match:	58.60%	Indels:	72
DB:	9	Gaps:	23

US-09-787-097-12 (1-1429) x AY442317 (1-4140)

```
Qy         28 GlyProHisTrpAspTrpAspValThrArgAlaGlyArgProGly---------------- 42
Db         18 GGCCCGCACTGG---------TACCCCGCAGCCAGCGGCCCCGGGGGTGTGGAGGGCTCG 68

Qy         43 LeuGlyAlaGlyLeuArgLeuProArgLeuLeuSerProProLeuArgProArgLeuLeu 62
Db         69 GCCGGCGGGCGGCGGCGGCGGGGGCGCCTCCTCCTGGCTGCTGGACGGGAACAGCTGGCT 128

Qy         63 LeuLeuLeuLeuLeuLeuProProProLeu-LeuLeuLeuLeuLeuProCysGluAlaGl 82
Db        129 GCTGTGCTATGGCTT---------CCTCTACCTGGCGCTCTAC----------------- 162

Qy         82 uAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGluCysAs 102
Db        163 ------------------------------GCGCAGGTGTCCCAGTCCAAGCCGTGCGA 191

Qy        102 pArg------ProCysValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCy 120
Db        192 GAGGACCGGCTCCTGCTTCTCG---GGCCGCTGTGTCAACTCCACC------TGCCTCTG 242

Qy        120 sProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGl 140
Db        243 CGACCCGGGCTGGGTGGGGGACCAGTGCCAGCACTGCCAGGGCAGGTTCAAGTTAACAGA 302

Qy        140 ySerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTr 160
Db        303 ACCTTCTGGATATTTAACAGATGGCCCAATTAACTATAAATATAAAACTAAATGTACTTG 362

Qy        160 pLeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGl 180
Db        363 GCTCATTGAAGGCTATCCAAATGCAGTGTTAAGATTAAGATTCAATCATTTTGCTACAGA 422

Qy        180 uCysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAl 200
Db        423 ATGTAGCTGGGATCATATGTATGTTTATGATGGAGATTCAATATATGCACCTTTAATAGC 482

Qy        200 aAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGluValVa 220
Db        483 TGTACTTAGTGGTTTGATAGTCCCTGAAATAAGGGGCAATGAAACTGTGCCTGAAGTTGT 542

Qy        220 lAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGl 240
Db        543 TACTACATCTGGCTATGCACTGTTACATTTTTTTAGTGATGCTGCGTATAATCTAACTGG 602

Qy        240 yPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCy 260
Db        603 TTTCAACATTTTCTATTCAATCAATTCTTGTCCTAACAATTGCTCTGGTCATGGGAAGTG 662

Qy        260 sLysIleSerAsnSer---SerGluThrValGluCysGluCysSerGluAsnTrpLysGl 279
Db        663 TACAACTAGTGTCTCTGTTCCAAGTCAAGTATATTGTGAATGTGATAAATACTGGAAGGG 722

Qy        279 yGluAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCy 299
Db        723 TGAAGCTTGTGATATTCCTTACTGTAAAGCCAATTGCGGCAGTCCAGATCACGGTTACTG 782

Qy        299 sAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSe 319
Db        783 TGACCTGACTGGAGAAAAATTATGTGTCTGCAATGATAGTTGGCAAGGTCCTGATTGTTC 842

Qy        319 rValProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeuLys-- 338
Db        843 TTTGAATGTTCCCTCTACTGAGTCTTACTGGATT---------CTGCCAAACGTTAAACC 893

Qy        339 ------------LeuProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTr 354
Db        894 CTTCAGTCCTTCTGTAGGTCGGGCTTCACATAAAGCAGTTTTACACGGGAAATTTATGTG 953

Qy        354 pValValGlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLe 374
Db        954 GGTGATTGGTGGATATACTTTTAACTACAGTTCTTTTCAAATGGTCCTAAATTACAATTT 1013

Qy        374 uAlaSerArgGluTrpLeuProLeuAsnArgSerValAsnAsnValValValArgTyrGl 394
Db       1014 AGAAAGCAGTATATGG---AATGTAGGAACTCCATCAAGGGGACCTCTCCAGAGATATGG 1070

Qy        394 yHisSerLeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLysIleAspProTh 414
Db       1071 ACACTCTCTTGCTTTATATCAGGAAAACATCTTTATGTATGGAGGCAGAATTGAAACAAA 1130

Qy        414 r---GlyAsnValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLe 433
Db       1131 TGATGGCAATGTCACAGATGAATTATGGGTTTTTAACATACATAGTCAGTCATGGAGTAC 1190

Qy        433 uLeuThrProLysAla---------LysGluGlnTyrAlaValValGlyHisSerAlaHi 450
Db       1191 AAAAACTCCTACTGTTCTTGGACATGGTCAGCAGTATGCTGTGGAGGGACATTCAGCACA 1250

Qy        450 sIleValThrLeuLysAsnGlyArgValValMetLeuValIlePheGlyHisCysProLe 470
Db       1251 TATTATGGAGTTGGATAGTAGAGATGTTGTCATGATCATAATATTTGGATATTCTGCAAT 1310

Qy        470 uTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLe 490
Db       1311 ATATGGTTATACAAGCAGCATACAGGAATACCATATCTCATCAAACACTTGGCTTGTTCC 1370

Qy        490 uHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisSerSerValTyrAspHisAr 510
Db       1371 AGAAACTAAAGGAGCTATTGTACAAGGTGGATATGGCCATACTAGTGTGTATGATGAAAT 1430

Qy        510 gThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLe 530
Db       1431 AACAAAGTCCATTTATGTTCATGGAGGGTATAAAGCATTGCCAGGGAACAAATATGGATT 1490

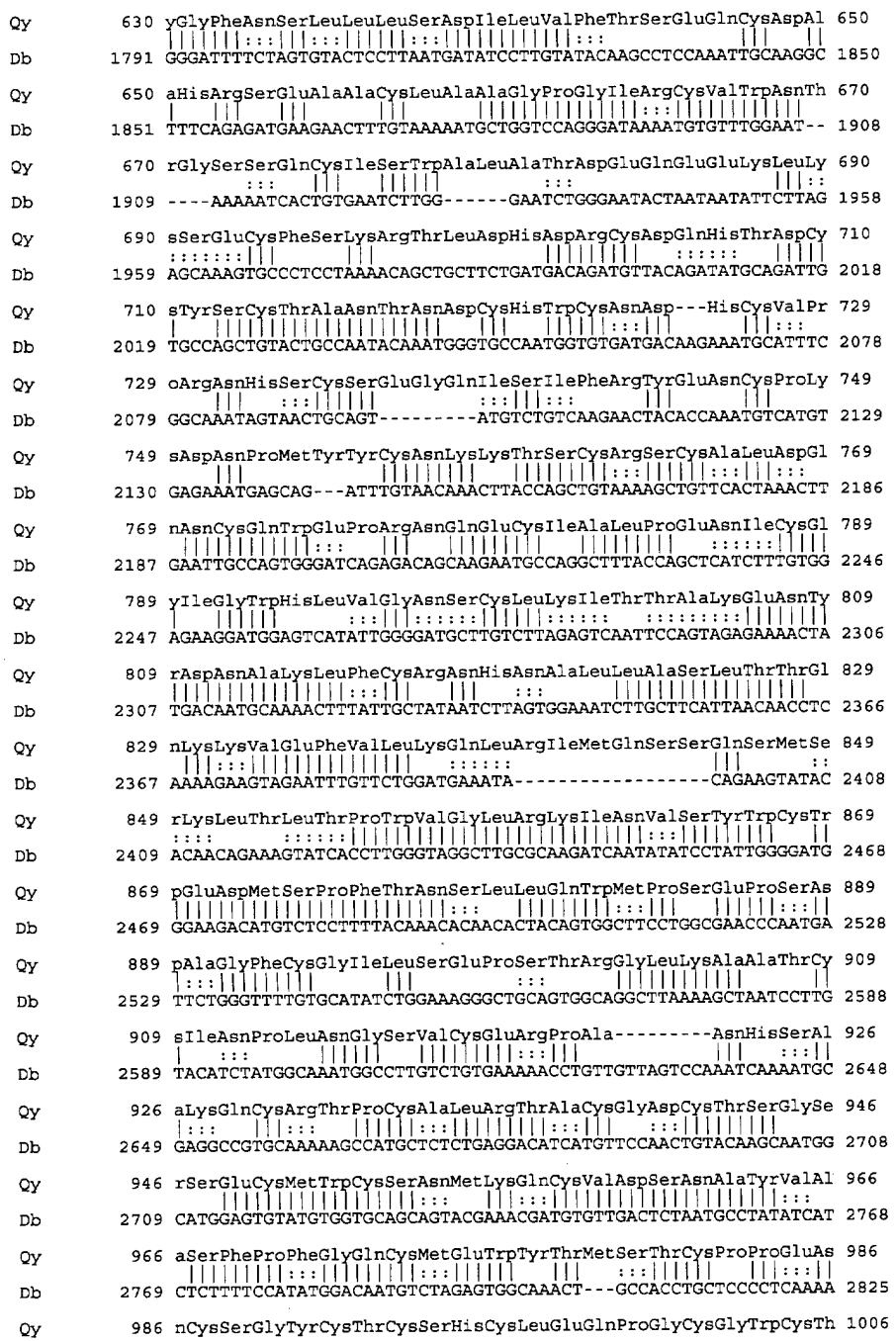
Qy        530 uAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetTrpThrIleLeuLysAspSe 550
Db       1491 GGTTGATGATCTTTATAAATATGAAGTTAACACTAAGACTTGGACTATTTTGAAAGAAAG 1550

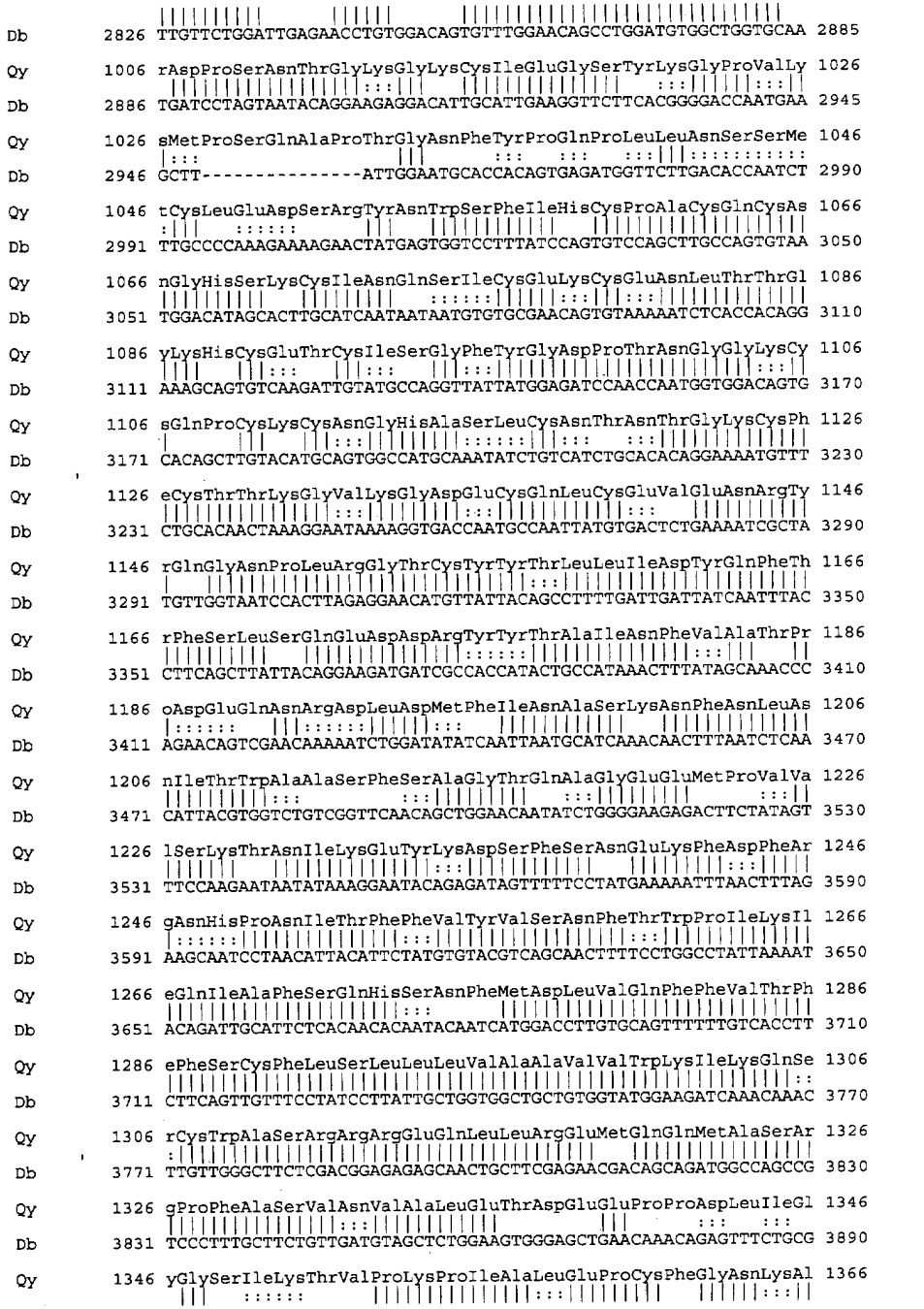
Qy        550 rArgPhePheArgTyrLeuHisThrAlaValIleValSerGlyThrMetLeuValPheGl 570
Db       1551 TGGGTTTGCCAGATACCTTCATTCAGCTGTTCTTATCAATGGAGCTATGCTTATTTTTGG 1610

Qy        570 yGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysPheSerSerAspPh 590
Db       1611 AGGAAATACCCATAATGACACTTCCTTGAGTAACGGTGCAAAATGTTTTTCTGCCGATTT 1670

Qy        590 eMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuProArgProAspLeuHisHi 610
Db       1671 CCTGGCATATGACATAGCTTGTGATGAATGGAAAATACTACCAAAACCAAATCTTCATAG 1730

Qy        610 sAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrMetTyrValPheGl 630
Db       1731 AGATGTCAACAGATTTGGACACTCTGCAGTAGTCATTAACGGGTCCATGTATATATTTGG 1790
```





```
Db       3891 AGGGCCATTAGAGGGGGCACCCAAGCCAATTGCCATTGAACCATGTGCTGGGAACAGAGC 3950

Qy       1366 aAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyGlyIleProProProGl 1386
              ||||||||||:::||||||:::  ||||||||||||     |||  ||||||||||||
Db       3951 TGCTGTTCTGACTGTGTTTCTTTGTCTACCACGAGGATCATCAGGTGCCCCTCCCCCTGG 4010

Qy       1386 yGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetProIleVa 1406
              ||||||||||||||||:::|||||||||||||:::|||||||||||||||
Db       4011 GCAGTCAGGCCTTGCAATTGCCAGTGCCCTAATAGATATTTCACAACAGAAAGCTTCAGA 4070

Qy       1406 lTyrLysGluLysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnProGl 1426
                 |||:::|||:::      ||||||||||||||||              |||   ||
Db       4071 TAGTAAAGATAAGACTTCTGGAGTCCGGAATCGAAAACACCTTTCAACACGTCAA---GG 4127

Qy       1426 yThrCysIle 1429
              |||||||:::
Db       4128 AACTTGTGTC 4137
```

RESULT 14
AK127277

LOCUS	AK127277 4463 bp mRNA linear PRI 09-SEP-2003
DEFINITION	Homo sapiens cDNA FLJ45344 fis, clone BRHIP3010289, moderately similar to Attractin precursor.
ACCESSION	AK127277
VERSION	AK127277.1 GI:34534122
KEYWORDS	oligo capping; fis (full insert sequence).
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1
AUTHORS	Oshima,A., Takahashi-Fujii,A., Tanase,T., Imose,N., Takeuchi,K., Arita,M., Musashino,K., Yuuki,H., Hara,H., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
TITLE	NEDO human cDNA sequencing project
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 4463)
AUTHORS	Isogai,T. and Yamamoto,J.
TITLE	Direct Submission
JOURNAL	Submitted (15-JUL-2003) Takao Isogai, FLJ Project(HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT	NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES Location/Qualifiers

```
     source          1. .4463
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="BRHIP3010289"
                     /tissue_type="hippocampus"
                     /clone_lib="BRHIP3"
                     /note="cloning vector: pME18SFL3"
     CDS             303. .4004
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAC86914.1"
                     /db_xref="GI:34534123"
                     /translation="MYVYDGDSIYAPLIAVLSGLIVPEIRGNETVPEVVTTSGYALLH
                     FFSDAAYNLTGFNIFYSINSCPNNCSGHGKCTTSVSVPSQVYCECDKYWKGEACDIPY
                     CKANCGSPDHGYCDLTGEKLCVCNDSWQGPDCSLNVPSTESYWILPNVKPFSPSVGRA
                     SHKAVLHGKFMWVIGGYTFNYSSFQMVLNYNLESSIWNVGTPSRGPLQRYGHSLALYQ
                     ENIFMYGGRIETNDGNVTDELWVFNIHSQSWSTKTPTVLGHGQQYAVEGHSAHIMELD
                     SRDVVMIIIFGYSAIYGYTSSIQEYHISSNTWLVPETKGAIVQGGYGHTSVYDEITKS
                     IYVHGGYKALPGNKYGLVDDLYKYEVNTKTWTILKESGFARYLHSAVLINGAMLIFGG
                     NTHNDTSLSNGAKCFSADFLAYDIACDEWKILPKPNLHRDVNRFGHSAVVINGSMYIF
                     GGFSSVLLNDILVYKPPNCKAFRDEELCKNAGPGIKCVWNKNHCESWESGNTNNILRA
                     KCPPKTAASDDRCYRYADCASCTANTNGCQWCDDKKCISANSNCSMSVKNYTKCHVRN
                     EQICNKLTSCKSCSLNLNCQWDQRQQECQALPAHLCGEGWSHIGDACLRVNSSRENYD
                     NAKLYCYNLSGNLASLTTSKEVEFVLDEIQKYTQQKVSPWVGLRKINISYWGWEDMSP
                     FTNTTLQWLPGEPNDSGFCAYLERAAVAGLKANPCTSMANGLVCEKPVVSPNQNARPC
                     KKPCSLRTSCSNCTSNGMECMWCSSTKRCVDSNAYIISFPYGQCLEWQTATCSPQNCS
                     GLRTCGQCLEQPGCGWCNDPSNTGRGHCIEGSSRGPMKLIGMHHNEMVLDTNLCPKEK
                     NYEWSFIQCPACQCNGHSTCINNNVCEQCKNLTTGKQCQDCMPGYYGDPTNGGQCTAC
                     TCSGHANICHLHTGKCFCTTKGIKGDQCQLCDSENRYVGNPLRGTCYYSLLIDYQFTF
                     SLLQEDDRHHTAINFIANPEQSNKNLDISINASNNFNLNITWSVGSTAGTISGEETSI
                     VSKNNIKEYRDSFSYEKFNFRSNPNITFYVYVSNFSWPIKIQIAFSQHNTIMDLVQFF
                     VTFFSCFLSLLLVAAVVWKIKQTCWASRRREQLLRERQQMASRPFASVDVALEVGAEQ
                     TEFLRGPLEGAPKPIAIEPCAGNRAAVLTVFLCLPRGSSGAPPPGQSGLAIASALIDI
                     SQQKASDSKDKTSGVRNRKHLSTRQGTCV"
```

ORIGIN

Alignment Scores:

Pred. No.:	4.54e-280	Length:	4463
Score:	4622.00	Matches:	823
Percent Similarity:	76.22%	Conservative:	206
Best Local Similarity:	60.96%	Mismatches:	277
Query Match:	58.40%	Indels:	44
DB:	9	Gaps:	19

US-09-787-097-12 (1-1429) x AK127277 (1-4463)

```
Qy         96 AlaGluAlaLysGluCysAspArg------ProCysValAsnGlyGlyArgCysAsnPro 113
              ::::::::::|||   |||:::|||          |||   :::   |||||||||
Db         36 TCCCAGTCCAAACCGTGCGAGAGGACCGGCTCCTGCTTCTCG---GGCCGCTGTGTCAAC 92

Qy        114 GlyThrGlyGlnCysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGly 133
                 |||      |||:::|||       ||||||||||||:::|||||||||||||||
Db         93 TCCACC------TGCCTCTGCGACCCGGGCTGGGTGGGGGACCAGTGCCAGCACTGCCAG 146

Qy        134 GlyArgPheArgLeuThrGlySerSerGlyPheValThrAspGlyProGlyAsnTyrLys 153
              |||||||||:::|||||||      ||||||::::::||||||||||||   |||||||||
Db        147 GGCAGGTTCAAGTTAACAGAACCTTCTGGATATTTAACAGATGGCCCAATTAACTATAAA 206

Qy        154 TyrLysThrLysCysThrTrpLeuIleGluGlyGlnProAsnArgIleMetArgLeuArg 173
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 TATAAAACTAAATGTACTTGGCTCATTGAAGGCTATCCAAATGCAGTGTTAAGATTAAGA 266

Qy        174 PheAsnHisPheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGlyAspSer 193
              |||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||||||
Db        267 TTCAATCATTTTGCTACAGAATGTAGCTGGGATCATATGTATGTTTATGATGGAGATTCA 326

Qy        194 IleTyrAlaProLeuValAlaAlaPheSerGlyLeuIleValProGluArgAspGlyAsn 213
              ||||||||||||||:::|||      |||||||||||||||||||||||      ||||||
Db        327 ATATATGCACCTTTAATAGCTGTACTTAGTGGTTTGATAGTCCCTGAAATAAGGGGCAAT 386

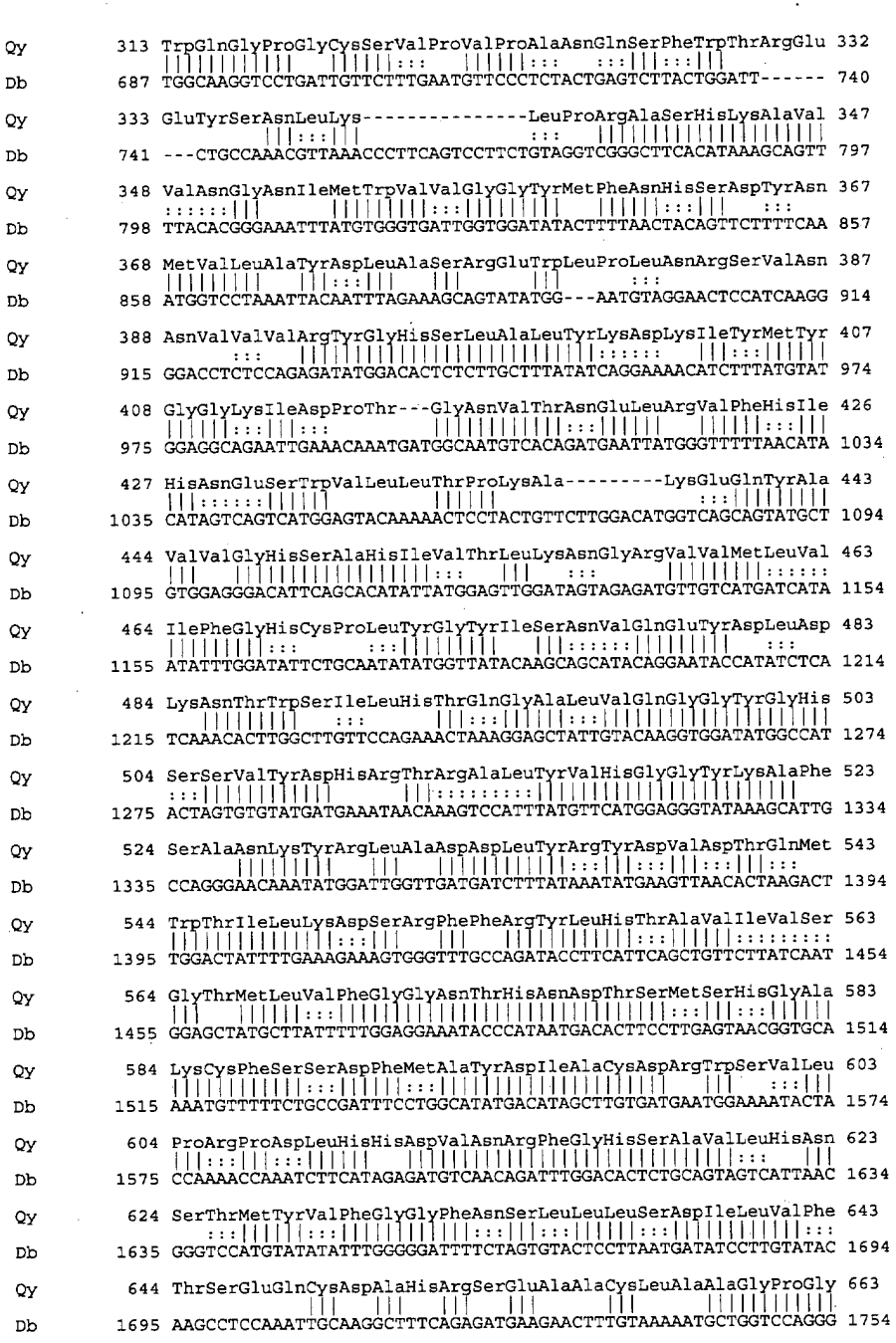
Qy        214 GluThrValProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAsp 233
              ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db        387 GAAACTGTGCCTGAAGTTGTTACTACATCTGGCTATGCACTGTTACATTTTTTTAGTGAT 446

Qy        234 AlaAlaTyrAsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsn 253
              ||||||||||||||||||||||||||||||  ||||||   :::   ||||||||||||
Db        447 GCTGCGTATAATCTAACTGGTTTCAACATTTTCTATTCAATCAATTCTTGTCCTAACAAT 506

Qy        254 CysSerGlyArgGlyGluCysLysIleSerAsnSer---SerGluThrValGluCysGlu 272
              |||||||||   |||:::|||      |||   |||                |||   ||||||
Db        507 TGCTCTGGTCATGGGAAGTGTACAACTAGTGTCTCTGTTCCAAGTCAAGTATATTGTGAA 566

Qy        273 CysSerGluAsnTrpLysGlyGluAlaCysAspIleProHisCysThrAspAsnCysGly 292
              |||   :::   ||||||||||||||||||||||||||||:::|||   ·   |||||||||
Db        567 TGTGATAAATACTGGAAGGGTGAAGCTTGTGATATTCCTTACTGTAAAGCCAATTGCGGC 626

Qy        293 PheProHisArgGlyIleCysAsnSerSerAspValArgGlyCysSerCysPheSerAsp 312
              |||      |||   |||:::   :::      :::   |||   |||
Db        627 AGTCCAGATCACGGTTACTGTGACCTGACTGGAGAAAAATTATGTGTCTGCAATGATAGT 686
```


Qy 313 TrpGlnGlyProGlyCysSerValProValProAlaAsnGlnSerPheTrpThrArgGlu 332
Db 687 TGGCAAGGTCCTGATTGTTCTTTGAATGTTCCCTCTACTGAGTCTTACTGGATT------ 740
Qy 333 GluTyrSerAsnLeuLys--------------LeuProArgAlaSerHisLysAlaVal 347
Db 741 ---CTGCCAAACGTTAAACCCTTCAGTCCTTCTGTAGGTCGGGCTTCACATAAAGCAGTT 797
Qy 348 ValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHisSerAspTyrAsn 367
Db 798 TTACACGGGAAATTTATGTGGGTGATTGGTGGATATACTTTTAACTACAGTTCTTTTCAA 857
Qy 368 MetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSerValAsn 387
Db 858 ATGGTCCTAAATTACAATTTAGAAAGCAGTATATGG---AATGTAGGAACTCCATCAAGG 914
Qy 388 AsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyrMetTyr 407
Db 915 GGACCTCTCCAGAGATATGGACACTCTCTTGCTTTATATCAGGAAAACATCTTTATGTAT 974
Qy 408 GlyGlyLysIleAspProThr---GlyAsnValThrAsnGluLeuArgValPheHisIle 426
Db 975 GGAGGCAGAATTGAAACAAATGATGGCAATGTCACAGATGAATTATGGGTTTTTAACATA 1034
Qy 427 HisAsnGluSerTrpValLeuLeuThrProLysAla---------LysGluGlnTyrAla 443
Db 1035 CATAGTCAGTCATGGAGTACAAAAACTCCTACTGTTCTTGGACATGGTCAGCAGTATGCT 1094
Qy 444 ValValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuVal 463
Db 1095 GTGGAGGGACATTCAGCACATATTATGGAGTTGGATAGTAGAGATGTTGTCATGATCATA 1154
Qy 464 IlePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAsp 483
Db 1155 ATATTTGGATATTCTGCAATATATGGTTATACAAGCAGCATACAGGAATACCATATCTCA 1214
Qy 484 LysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHis 503
Db 1215 TCAAACACTTGGCTTGTTCCAGAAACTAAAGGAGCTATTGTACAAGGTGGATATGGCCAT 1274
Qy 504 SerSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPhe 523
Db 1275 ACTAGTGTGTATGATGAAATAACAAAGTCCATTTATGTTCATGGAGGGTATAAAGCATTG 1334
Qy 524 SerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMet 543
Db 1335 CCAGGGAACAAATATGGATTGGTTGATGATCTTTATAAATATGAAGTTAACACTAAGACT 1394
Qy 544 TrpThrIleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSer 563
Db 1395 TGGACTATTTTGAAAGAAAGTGGGTTTGCCAGATACCTTCATTCAGCTGTTCTTATCAAT 1454
Qy 564 GlyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAla 583
Db 1455 GGAGCTATGCTTATTTTTGGAGGAAATACCCATAATGACACTTCCTTGAGTAACGGTGCA 1514
Qy 584 LysCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeu 603
Db 1515 AAATGTTTTTCTGCCGATTTCCTGGCATATGACATAGCTTGTGATGAATGGAAAATACTA 1574
Qy 604 ProArgProAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsn 623
Db 1575 CCAAAACCAAATCTTCATAGAGATGTCAACAGATTTGGACACTCTGCAGTAGTCATTAAC 1634
Qy 624 SerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPhe 643
Db 1635 GGGTCCATGTATATATTTGGGGGATTTTCTAGTGTACTCCTTAATGATATCCTTGTATAC 1694
Qy 644 ThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGly 663
Db 1695 AAGCCTCCAAATTGCAAGGCTTTCAGAGATGAAGAACTTTGTAAAAATGCTGGTCCAGGG 1754
Qy 664 IleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAsp 683
Db 1755 ATAAAATGTGTTTGGAAT------AAAAATCACTGTGAATCTTGG------GAATCTGGG 1802
Qy 684 GluGlnGluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArg 703
Db 1803 AATACTAATAATATTCTTAGAGCAAAGTGCCCTCCTAAAACAGCTGCTTCTGATGACAGA 1862
Qy 704 CysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCys 723
Db 1863 TGTTACAGATATGCAGATTGTGCCAGCTGTACTGCCAATACAAATGGGTGCCAATGGTGT 1922
Qy 724 AsnAsp---HisCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePhe 742
Db 1923 GATGACAAGAAATGCATTTCGGCAAATAGTAACTGCAGT---------ATGTCTGTCAAG 1973
Qy 743 ArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCys 762
Db 1974 AACTACACCAAATGTCATGTGAGAAATGAGCAG---ATTTGTAACAAACTTACCAGCTGT 2030
Qy 763 ArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAla 782
Db 2031 AAAAGCTGTTCACTAAACTTGAATTGCCAGTGGGATCAGAGACAGCAAGAATGCCAGGCT 2090
Qy 783 LeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIle 802
Db 2091 TTACCAGCTCATCTTTGTGGAGAAGGATGGAGTCATATTGGGGATGCTTGTCTTAGAGTC 2150
Qy 803 ThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeu 822
Db 2151 AATTCCAGTAGAGAAAACTATGACAATGCAAAACTTTATTGCTATAATCTTAGTGGAAAT 2210
Qy 823 LeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMet 842
Db 2211 CTTGCTTCATTAACAACCTCAAAAGAAGTAGAATTTGTTCTGGATGAAATA--------- 2261
Qy 843 GlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIle 862
Db 2262 ---------CAGAAGTATACACAACAGAAAGTATCACCTTGGGTAGGCTTGCGCAAGATC 2312
Qy 863 AsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrp 882
Db 2313 AATATATCCTATTGGGGATGGGAAGACATGTCTCCTTTTACAAACACAACACTACAGTGG 2372
Qy 883 MetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArg 902
Db 2373 CTTCCTGGCGAACCCAATGATTCTGGGTTTTGTGCATATCTGGAAAGGGCTGCAGTGGCA 2432
Qy 903 GlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAla 922
Db 2433 GGCTTAAAAGCTAATCCTTGTACATCTATGGCAAATGGCCTTGTCTGTGAAAAACCTGTT 2492
Qy 923 ---------AsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCys 939
Db 2493 GTTAGTCCAAATCAAAATGCGAGGCCGTGCAAAAAGCCATGCTCTCTGAGGACATCATGT 2552
Qy 940 GlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysVal 959
Db 2553 TCCAACTGTACAAGCAATGGCATGGAGTGTATGTGGTGCAGCAGTACGAAACGATGTGTT 2612
Qy 960 AspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMet 979
Db 2613 GACTCTAATGCCTATATCATCTCTTTTCCATATGGACAATGTCTAGAGTGGCAAACT--- 2669
Qy 980 SerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGln 999
Db 2670 GCCACCTGCTCCCCTCAAAATTGTTCTGGATTGAGAACCTGTGGACAGTGTTTGGAACAG 2729
Qy 1000 ProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGly 1019
Db 2730 CCTGGATGTGGCTGGTGCAATGATCCTAGTAATACAGGAAGAGGACATTGCATTGAAGGT 2789
Qy 1020 SerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGln 1039

```
Db       2790 TCTTCACGGGGACCAATGAAGCTT----------------ATTGGAATGCACCACAATGAG 2834

Qy       1040 ProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHis 1059

Db       2835 ATGGTTCTTGACACCAATCTTTGCCCCAAAGAAAAGAACTATGAGTGGTCCTTTATCCAG 2894

Qy       1060 CysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLys 1079

Db       2895 TGTCCAGCTTGCCAGTGTAATGGACATAGCACTTGCATCAATAATAATGTGTGCGAACAG 2954

Qy       1080 CysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAsp 1099

Db       2955 TGTAAAAATCTCACCACAGGAAAGCAGTGTCAAGATTGTATGCCAGGTTATTATGGAGAT 3014

Qy       1100 ProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsn 1119

Db       3015 CCAACCAATGGTGGACAGTGCACAGCTTGTACATGCAGTGGCCATGCAAATATCTGTCAT 3074

Qy       1120 ThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeu 1139

Db       3075 CTGCACACAGGAAAATGTTTCTGCACAACTAAAGGAATAAAAGGTGACCAATGCCAATTA 3134

Qy       1140 CysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeu 1159

Db       3135 TGTGACTCTGAAAATCGCTATGTTGGTAATCCACTTAGAGGAACATGTTATTACAGCCTT 3194

Qy       1160 LeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAla 1179

Db       3195 TTGATTGATTATCAATTTACCTTCAGCTTATTACAGGAAGATGATCGCCACCATACTGCC 3254

Qy       1180 IleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAla 1199

Db       3255 ATAAACTTTATAGCAAACCCAGAACAGTCGAACAAAAATCTGGATATATCAATTAATGCA 3314

Qy       1200 SerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAla 1219

Db       3315 TCAAACAACTTTAATCTCAACATTACGTGGTCTGTCGGTTCAACAGCTGGAACAATATCT 3374

Qy       1220 GlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSer 1239

Db       3375 GGGGAAGAGACTTCTATAGTTTCCAAGAATAATATAAAGGAATACAGAGATAGTTTTTCC 3434

Qy       1240 AsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsn 1259

Db       3435 TATGAAAAATTTAACTTTAGAAGCAATCCTAACATTACGTTCTATGTGTACGTCAGCAAC 3494

Qy       1260 PheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeu 1279

Db       3495 TTTTCCTGGCCTATTAAAATACAGATTGCATTCTCACAACACAATACAATCATGGACCTT 3554

Qy       1280 ValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuLeuValAlaAlaVal 1299

Db       3555 GTGCAGTTTTTTGTCACCTTCTTCAGTTGTTTCCTATCCTTATTGCTGGTGGCTGCTGTG 3614

Qy       1300 ValTrpLysIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeuLeuArgGlu 1319

Db       3615 GTATGGAAGATCAAACAAACTTGTTGGGCTTCTCGACGGAGAGAGCAACTTCTTCGAGAA 3674

Qy       1320 MetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThrAspGlu 1339

Db       3675 CGACAGCAGATGGCCAGCCGTCCCTTTGCTTCTGTTGATGTAGCTCTGGAAGTGGGAGCT 3734

Qy       1340 GluProProAspLeuIleGlyGlySerIleLysThrValProLysProIleAlaLeuGlu 1359

Db       3735 GAACAAACAGAGTTTCTGCGAGGGCCATTAGAGGGGGCACCCAAGCCAATTGCCATTGAA 3794

Qy       1360 ProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProArgGlyLeu 1379

Db       3795 CCATGTGCTGGGAACAGAGCTGCTGTTCTGACTGTGTTTCTTTGTCTACCACGAGGATCA 3854

Qy       1380 GlyGlyIleProProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIle 1399

Db       3855 TCAGGTGCCCCTCCCCCTGGGCAGTCAGGCCTTGCAATTGCCAGTGCCCTAATAGATATT 3914

Qy       1400 SerGlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAsnArgLysGln 1419

Db       3915 TCACAACAGAAAGCTTCAGATAGTAAAGATAAGACTTCTGGAGTCCGGAATCGAAAACAC 3974

Qy       1420 GlnProProAlaGlnProGlyThrCysIle 1429

Db       3975 CTTTCAACACGTCAA---GGAACTTGTGTC 4001
```

RESULT 15
AR164818
LOCUS AR164818 2625 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 18 from patent US 6274339.
ACCESSION AR164818
VERSION AR164818.1 GI:16238029
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 2625)
AUTHORS Moore,K. and Nagle,D.Lynn.
TITLE Methods and compositions for the diagnosis and treatment of body weight disorders, including obesity
JOURNAL Patent: US 6274339-A 18 14-AUG-2001;
FEATURES Location/Qualifiers
source 1. .2625
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN

Alignment Scores:

Pred. No.:	6.29e-250	Length:	2625
Score:	4136.50	Matches:	759
Percent Similarity:	95.48%	Conservative:	2
Best Local Similarity:	95.23%	Mismatches:	22
Query Match:	52.26%	Indels:	14
DB:	6	Gaps:	3

US-09-787-097-12 (1-1429) x AR164818 (1-2625)

```
Qy          2 ValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAlaAla 21

Db         10 GTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCG 69

Qy         22 LeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArgPro 41

Db         70 CTCGCGGGCAGGAGCGGCGGGCCGCACCGACCCTGCACCGCGACAGGGGCCTGGAGGCCG 129

Qy         42 GlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSer----------------ProPro 56

Db        130 GGACCGCGCGCCCGGCTGTGTCTCCCGCGGGTGCTGTCGCGGGCGCTGCCCCCGCCGCCG 189

Qy         57 LeuArgProArgLeuLeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeu 76

Db        190 CTGCTGCCG---CTGCTCTTTTCGCTGCTGCTGCTGCCG---------------------- 225

Qy         77 LeuProCysGluAlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAla 96

Db        226 CTGCCCCGGGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCC 285

Qy         97 GluAlaLysGluCysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGly 116

Db        286 GAGGCCAAGGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGC 345

Qy        117 GlnCysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPhe 136

Db        346 CAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTC 405

Qy        137 ArgLeuThrGlySerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThr 156

Db        406 AGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACG 465
```

```
Qy      157 LysCysThrTrpLeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHis 176
Db      466 AAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCAT 525

Qy      177 PheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAla 196
Db      526 TTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCA 585

Qy      197 ProLeuValAlaAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrVal 216
Db      586 CCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTC 645

Qy      217 ProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyr 236
Db      646 CCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTAT 705

Qy      237 AsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGly 256
Db      706 AATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGC 765

Qy      257 ArgGlyGluCysLysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsn 276
Db      766 CGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAAC 825

Qy      277 TrpLysGlyGluAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArg 296
Db      826 TGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGA 885

Qy      297 GlyIleCysAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyPro 316
Db      886 GGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCT 945

Qy      317 GlyCysSerValProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsn 336
Db      946 GGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAAC 1005

Qy      337 LeuLysLeuProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValVal 356
Db     1006 TTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTT 1065

Qy      357 GlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSer 376
Db     1066 GGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCT 1125

Qy      377 ArgGluTrpLeuProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSer 396
Db     1126 AGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCT 1185

Qy      397 LeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsn 416
Db     1186 TTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAAT 1245

Qy      417 ValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrPro 436
Db     1246 GTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCT 1305

Qy      437 LysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsn 456
Db     1306 AAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAAT 1365

Qy      457 GlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsn 476
Db     1366 GGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAAT 1425

Qy      477 ValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeu 496
Db     1426 GTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTT 1485

Qy      497 ValGlnGlyGlyTyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrVal 516
Db     1486 GTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTT 1545

Qy      517 HisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArg 536
Db     1546 CATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGA 1605

Qy      537 TyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeu 556
Db     1606 TATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTG 1665

Qy      557 HisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAsp 576
Db     1666 CACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGAC 1725

Qy      577 ThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAla 596
Db     1726 ACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCC 1785

Qy      597 CysAspArgTrpSerValLeuProArgProAspLeuHisHisAspValAsnArgPheGly 616
Db     1786 TGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGC 1845

Qy      617 HisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeu 636
Db     1846 CATTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTC 1905

Qy      637 LeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAla 656
Db     1906 CTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCT 1965

Qy      657 CysLeuAlaAlaGlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIle 676
Db     1966 TGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATC 2025

Qy      677 SerTrpAlaLeuAlaThrAspGluGlnGluGluLysLeuLysSerGluCysPheSerLys 696
Db     2026 TCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAA 2085

Qy      697 ArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsn 716
Db     2086 AGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAAC 2145

Qy      717 ThrAsnAspCysHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGlu 736
Db     2146 ACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAA 2205

Qy      737 GlyGlnIleSerIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCys 756
Db     2206 GGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGT 2265

Qy      757 AsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArg 776
Db     2266 AACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGG 2325

Qy      777 AsnGlnGluCysIleAlaLeuProGluAsnIleCys-GlyIleGlyTrp 792
Db     2326 AATCAGGAGTGCATTGCCCTGCCCGGTAGGCCTTGCAGGGTCATCTTGG 2374
```

Search completed: March 1, 2004, 05:12:59
Job time : 11835 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 1, 2004, 00:45:20 ; Search time 6892 Seconds
(without alignments)
6191.680 Million cell updates/sec

Title: US-09-787-097-12
Perfect score: 7915
Sequence: 1 MVAAAAATEARLRRRTAATA.....KSGAVRNRKQQPPAQPGTCI 1429

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09787097/runat_26022004_062935_23428/app_query.fasta_1.1607
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09787097_@CGN_1_1_4337_@runat_26022004_062935_23428 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	6217	78.5	3976	29	AY418587	AY418587 Homo sapi
	2	5980	75.6	3976	29	AY418589	AY418589 Mus muscu
	3	5018	63.4	3671	29	AY418588	AY418588 Pan trogl
	4	1746.5	22.1	1201	13	BX440935	BX440935 BX440935
	5	1619	20.5	964	13	BU506373	BU506373 AGENCOURT
	6	1491	18.8	840	14	CB520374	CB520374 UI-M-GI0-
	7	1421.5	18.0	933	12	BG678679	BG678679 602624477
	8	1402	17.7	1545	11	AK050882	AK050882 Mus muscu
	9	1397	17.7	770	13	BY761276	BY761276 BY761276
	10	1350	17.1	735	13	BY761390	BY761390 BY761390
	11	1315	16.6	777	14	CF745736	CF745736 UI-M-GV0-
	12	1310	16.6	681	12	BM783739	BM783739 K-EST0061
	13	1305	16.5	680	12	BM783258	BM783258 K-EST0061
	14	1305	16.5	680	12	BM783773	BM783773 K-EST0061
	15	1301	16.4	683	12	BM783788	BM783788 K-EST0061
	16	1296	16.4	677	12	BM783718	BM783718 K-EST0061
	17	1292	16.3	673	12	BM783685	BM783685 K-EST0061
	18	1291	16.3	1836	11	BC035157	BC035157 Homo sapi
	19	1290	16.3	677	12	BM783821	BM783821 K-EST0061
c	20	1254	15.8	722	13	BX669806	BX669806 BX669806
	21	1233	15.6	792	14	CD644584	CD644584 AGENCOURT
	22	1233	15.6	962	13	BX431593	BX431593 BX431593
c	23	1231.5	15.6	714	10	BE650220	BE650220 UI-M-BH2.
	24	1226	15.5	674	12	BG563729	BG563729 602584512
	25	1225	15.5	917	10	BF672370	BF672370 602150605
	26	1208.5	15.3	940	13	BQ956223	BQ956223 AGENCOURT
	27	1206	15.2	632	12	BM773003	BM773003 K-EST0057
	28	1206	15.2	632	12	BM783264	BM783264 K-EST0061
	29	1191	15.0	716	14	CF745901	CF745901 UI-M-GV0-
	30	1191	15.0	739	9	AL629411	AL629411 AL629411
	31	1187	15.0	706	12	BI089332	BI089332 602853292
	32	1183	14.9	809	13	BU286532	BU286532 603605969
	33	1179	14.9	663	14	CF745835	CF745835 UI-M-GV0-
	34	1163	14.7	743	14	CB519923	CB519923 UI-M-GI0-
	35	1161	14.7	633	10	BE894518	BE894518 601433126
	36	1130	14.3	626	10	BE980188	BE980188 UI-M-BG2-
	37	1126	14.2	953	10	BF313142	BF313142 601897661
	38	1125.5	14.2	1734	11	BC029592	BC029592 Homo sapi
	39	1122	14.2	591	12	BI850289	BI850289 imageqc_1
	40	1115.5	14.1	902	13	BU151781	BU151781 AGENCOURT
	41	1112	14.0	710	13	BU377616	BU377616 603811614
	42	1110	14.0	991	10	BE880905	BE880905 601490448
	43	1108	14.0	780	13	BU378893	BU378893 603813635
	44	1088	13.7	653	14	CB446934	CB446934 700912 MA
	45	1079	13.6	641	10	BB664616	BB664616 BB664616

ALIGNMENTS

RESULT 1
AY418587
LOCUS AY418587 3976 bp DNA linear GSS 17-DEC-2003
DEFINITION Homo sapiens ATRN gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION AY418587
VERSION AY418587.1 GI:39774547
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3976)

```
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 3976)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     These sequences were made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .3976
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     gene            <1. .>3976
                     /gene="ATRN"
                     /locus_tag="HCM6607"
ORIGIN

Alignment Scores:
Pred. No.:                0              Length:          3976
Score:                    6217.00        Matches:         1139
Percent Similarity:       86.10%         Conservative:    1
Best Local Similarity:    86.03%         Mismatches:      184
Query Match:              78.55%         Indels:          0
DB:                       29             Gaps:            0

US-09-787-097-12 (1-1429) x AY418587 (1-3976)

Qy        106 ValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTrpVal 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 GTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTG 61

Qy        126 GlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPheVal 145
              ||||||||||||||||||||||||||||||||
Db         62 GGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGNNNNNNNNNNNNNNNNNNNNNNNN 121

Qy        146 ThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeuIleGluGlyGln 165

Db        122 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNG 181

Qy        166 ProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHis 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 CCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCAT 241

Qy        186 LeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaAlaPheSerGlyLeu 205
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 TTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGNNNNNNN 301

Qy        206 IleValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGlyTyr 225

Db        302 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 361

Qy        226 AlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThrTyr 245

Db        362 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 421

Qy        246 SerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSer 265
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 NNTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAGC 481

Qy        266 SerGluThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIlePro 285
              |||:::||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 AGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCT 541

Qy        286 HisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspValArg 305
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 CACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGA 601

Qy        306 GlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAlaAsn 325
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 GGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAAC 661

Qy        326 GlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHisLys 345
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 CAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAAA 721

Qy        346 AlaValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHisSerAsp 365
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 GCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGAT 781

Qy        366 TyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSer 385
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        782 TATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCT 841

Qy        386 ValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyr 405
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        842 GTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATACAAGNNNNNNNNNNNN 901

Qy        406 MetTyrGlyGlyLysIleAspProThrGlyAsnValThrAsnGluLeuArgValPheHis 425

Db        902 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 961

Qy        426 IleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaValVal 445

Db        962 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1021

Qy        446 GlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValIlePhe 465

Db       1022 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1081

Qy        466 GlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsn 485
                                                                  ||||||
Db       1082 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNATAAGAAC 1141

Qy        486 ThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisSerSer 505
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1142 ACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGT 1201

Qy        506 ValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAla 525
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1202 GTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCC 1261

Qy        526 AsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetTrpThr 545
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1262 AATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACC 1321

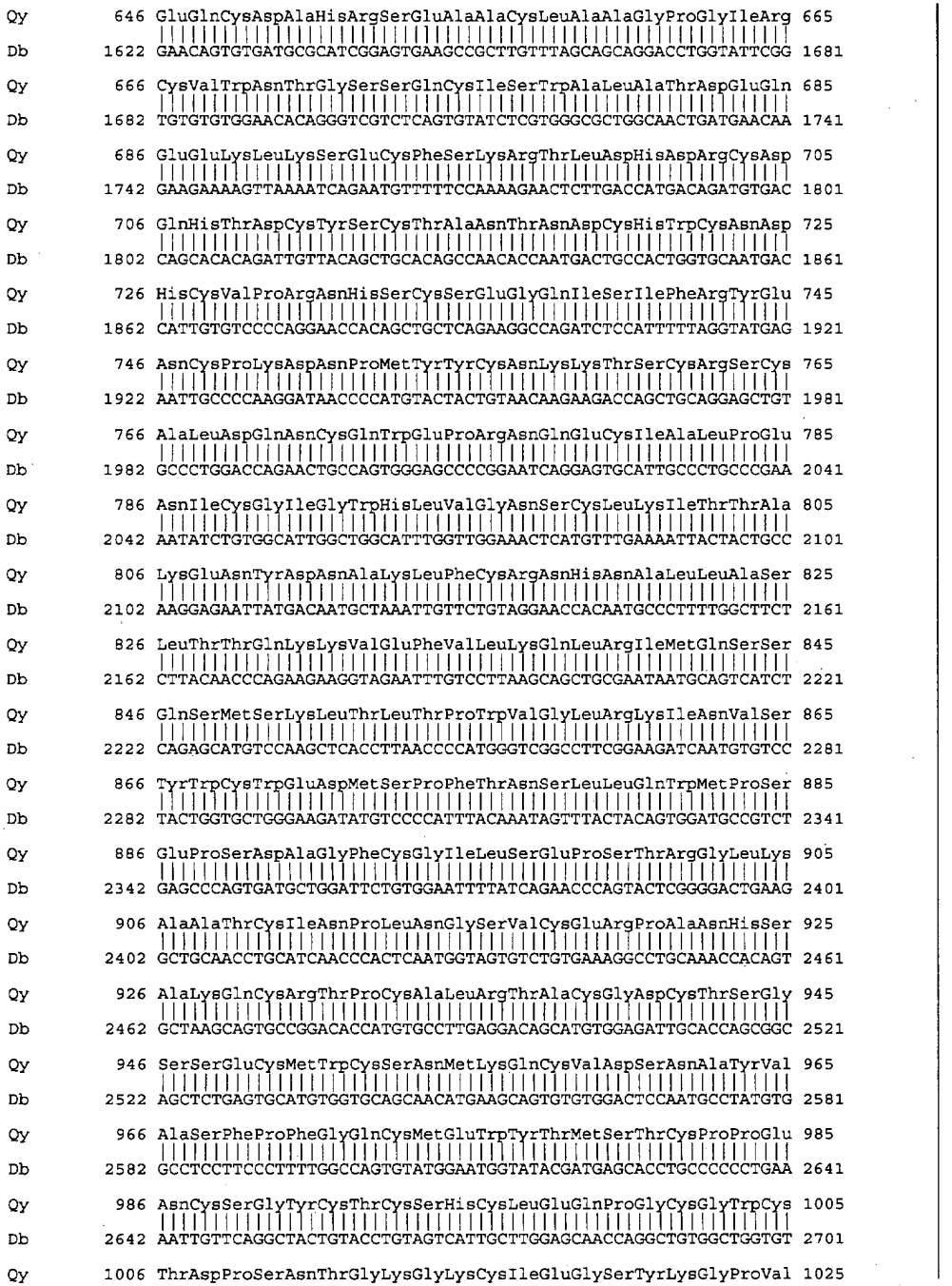
Qy        546 IleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerGlyThr 565
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1322 ATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACC 1381

Qy        566 MetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCys 585
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1382 ATGCTGGTGTTTGGAGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATGC 1441

Qy        586 PheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuProArg 605
              ||||||||||||||||||||||||||
Db       1442 TTCTCTTCAGATTTCATGGCCTATGACATTGNNNNNNNNNNNNNNNNNNNNNNNNNNN 1501

Qy        606 ProAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThr 625
                                                                  |||
Db       1502 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCACC 1561

Qy        626 MetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSer 645
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1562 ATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCG 1621
```


Qy 646 GluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArg 665
Db 1622 GAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGG 1681
Qy 666 CysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGln 685
Db 1682 TGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAA 1741
Qy 686 GluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArgCysAsp 705
Db 1742 GAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGAC 1801
Qy 706 GlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAsp 725
Db 1802 CAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGCAATGAC 1861
Qy 726 HisCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGlu 745
Db 1862 CATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAG 1921
Qy 746 AsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArgSerCys 765
Db 1922 AATTGCCCCAAGGATAACCCCATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGT 1981
Qy 766 AlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuProGlu 785
Db 1982 GCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAA 2041
Qy 786 AsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThrThrAla 805
Db 2042 AATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCC 2101
Qy 806 LysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSer 825
Db 2102 AAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCT 2161
Qy 826 LeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGlnSerSer 845
Db 2162 CTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCT 2221
Qy 846 GlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSer 865
Db 2222 CAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCC 2281
Qy 866 TyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSer 885
Db 2282 TACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCT 2341
Qy 886 GluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLys 905
Db 2342 GAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAG 2401
Qy 906 AlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSer 925
Db 2402 GCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGT 2461
Qy 926 AlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGly 945
Db 2462 GCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGGC 2521
Qy 946 SerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrVal 965
Db 2522 AGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGTG 2581
Qy 966 AlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSerThrCysProProGlu 985
Db 2582 GCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGAA 2641
Qy 986 AsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCys 1005
Db 2642 AATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGT 2701
Qy 1006 ThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyProVal 1025
Db 2702 ACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTG 2761
Qy 1026 LysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerSer 1045
Db 2762 AAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGC 2821
Qy 1046 MetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnCys 1065
Db 2822 ATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATGC 2881
Qy 1066 AsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrThr 1085
Db 2882 AACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCACA 2941
Qy 1086 GlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyGlyLys 1105
Db 2942 GGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAA 3001
Qy 1106 CysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCys 1125
Db 3002 TGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGC 3061
Qy 1126 PheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGluAsnArg 1145
Db 3062 TTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGA 3121
Qy 1146 TyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrGlnPhe 1165
Db 3122 TACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTTC 3181
Qy 1166 ThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIleAsnPheValAlaThr 1185
Db 3182 ACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACT 3241
Qy 1186 ProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeu 1205
Db 3242 CCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCTC 3301
Qy 1206 AsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProVal 1225
Db 3302 AACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTT 3361
Qy 1226 ValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPheAspPhe 1245
Db 3362 GTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTT 3421
Qy 1246 ArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpProIleLys 1265
Db 3422 CGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAAA 3481
Qy 1266 IleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePheValThr 1285
Db 3482 ATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCTGGTACAGTTCTTCGTGACT 3541
Qy 1286 PhePheSerCysPheLeuSerLeuLeuLeuValAlaAlaValValTrpLysIleLysGln 1305
Db 3542 TTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGTGGTTTGGAAGATCAAACAA 3601
Qy 1306 SerCysTrpAlaSerArgArgArgGluGlnLeuLeuArgGluMetGlnGlnMetAlaSer 1325
Db 3602 AGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGAGATGCAACAGATGGCCAGC 3661
Qy 1326 ArgProPheAlaSerValAsnValAlaLeuGluThrAspGluGluProProAspLeuIle 1345
Db 3662 CGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTATT 3721
Qy 1346 GlyGlySerIleLysThrValProLysProIleAlaLeuGluProCysPheGlyAsnLys 1365
Db 3722 GGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGAGCCGTGTTTTGGCAACAAA 3781
Qy 1366 AlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyGlyIleProProPro 1385

```
Db       3782 GCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCTGGGTGGCATCCCTCCTCCT 3841

Qy       1386 GlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetProIle 1405
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3842 GGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGATA 3901

Qy       1406 ValTyrLysGluLysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnPro 1425
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3902 GTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGCCT 3961

Qy       1426 GlyThrCysIle 1429
              ||||||||||||
Db       3962 GGGACCTGCATC 3973

RESULT 2
AY418589
LOCUS       AY418589               3976 bp    DNA     linear   GSS 17-DEC-2003
DEFINITION  Mus musculus ATRN gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY418589
VERSION     AY418589.1  GI:39774549
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 3976)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 3976)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     These sequences were made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .3976
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
     gene            <1. .>3976
                     /gene="ATRN"
                     /locus_tag="HCM6607"
ORIGIN

Alignment Scores:
Pred. No.:              0           Length:         3976
Score:                  5980.00     Matches:        1089
Percent Similarity:     84.14%      Conservative:   25
Best Local Similarity:  82.25%      Mismatches:     210
Query Match:            75.55%      Indels:         0
DB:                     29          Gaps:           0

US-09-787-097-12 (1-1429) x AY418589 (1-3976)

Qy        106 ValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTrpVal 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db          2 GTCAACGGCGGCCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCACGGGCTGGGTG 61

Qy        126 GlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPheVal 145
              ||||||||||||||||||||||||||||||||
Db         62 GGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGNNNNNNNNNNNNNNNNNNNNNNNN 121

Qy        146 ThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeuIleGluGlyGln 165

Db        122 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNG 181

Qy        166 ProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHis 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 CCAAATAGAATAATGAGACTTCGCTTCAACCATTTTGCTACAGAATGTAGCTGGGACCAT 241

Qy        186 LeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaAlaPheSerGlyLeu 205
              ||||||||||||||||||||||||||||||||||||:::|||||||||
Db        242 TTATATGTTTATGATGGGGACTCAATCTACGCACCTCTGATTGCTGCCTTTAGNNNNNNN 301

Qy        206 IleValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGlyTyr 225

Db        302 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 361

Qy        226 AlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThrTyr 245

Db        362 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 421

Qy        246 SerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSer 265
              ||||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db        422 NNTTTTGACATGTGTCCGAATAATTGCTCAGGCCGAGGAGAGTGTAAGAGCAGTAACAGC 481

Qy        266 SerGluThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIlePro 285
              |||     ||||||||||||||||||||||||||||||||:::|||||||||||||||
Db        482 AGCAGCGCTGTTGAGTGTGAATGTTCTGAAAACTGGAAAGGGGAGTCGTGTGACATTCCT 541

Qy        286 HisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspValArg 305
              ||||||||||||||||||||||||||||||||||||||||:::|||||||   |||
Db        542 CACTGTACAGACAACTGTGGCTTTCCTCACCGAGGCATCTGTAATGCAAGCGATACCAGA 601

Qy        306 GlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAlaAsn 325
              ||||||||||||||||||||||||||||||||||||||:::|||||||||||||||||||
Db        602 GGGTGCTCCTGCTTTCCTCACTGGCAGGGTCCTGGATGCTCAATTCCTGTGCCAGCTAAC 661

Qy        326 GlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHisLys 345
              ||||||||||||||||||||||||||||||:::|||||||||||||||||||||||||||
Db        662 CAGTCTTTTTGGACTCGAGAAGAATATTCTGATTTAAAGCTTCCCAGAGCCTCTCATAAA 721

Qy        346 AlaValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHisSerAsp 365
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 GCTGTGGTCAATGGAAATATAATGTGGGTTGTTGGCGGATATATGTTCAACCATTCAGAT 781

Qy        366 TyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSer 385
              |||:::|||||||||||||||||||||   ||||||||||||||||||||||   |||
Db        782 TACAGCATGGTTTTAGCGTATGACCTGACTTCTAGGGAATGGCTTCCACTAAACCATTCT 841

Qy        386 ValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyr 405
              |||||:::|||||||||||||||||||||||||||||||:::|||
Db        842 GTGAACAGTGTGGTTGTAAGATATGGTCATTCTTTGGCATTACATAAGNNNNNNNNNNNN 901

Qy        406 MetTyrGlyGlyLysIleAspProThrGlyAsnValThrAsnGluLeuArgValPheHis 425

Db        902 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 961

Qy        426 IleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaValVal 445

Db        962 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1021

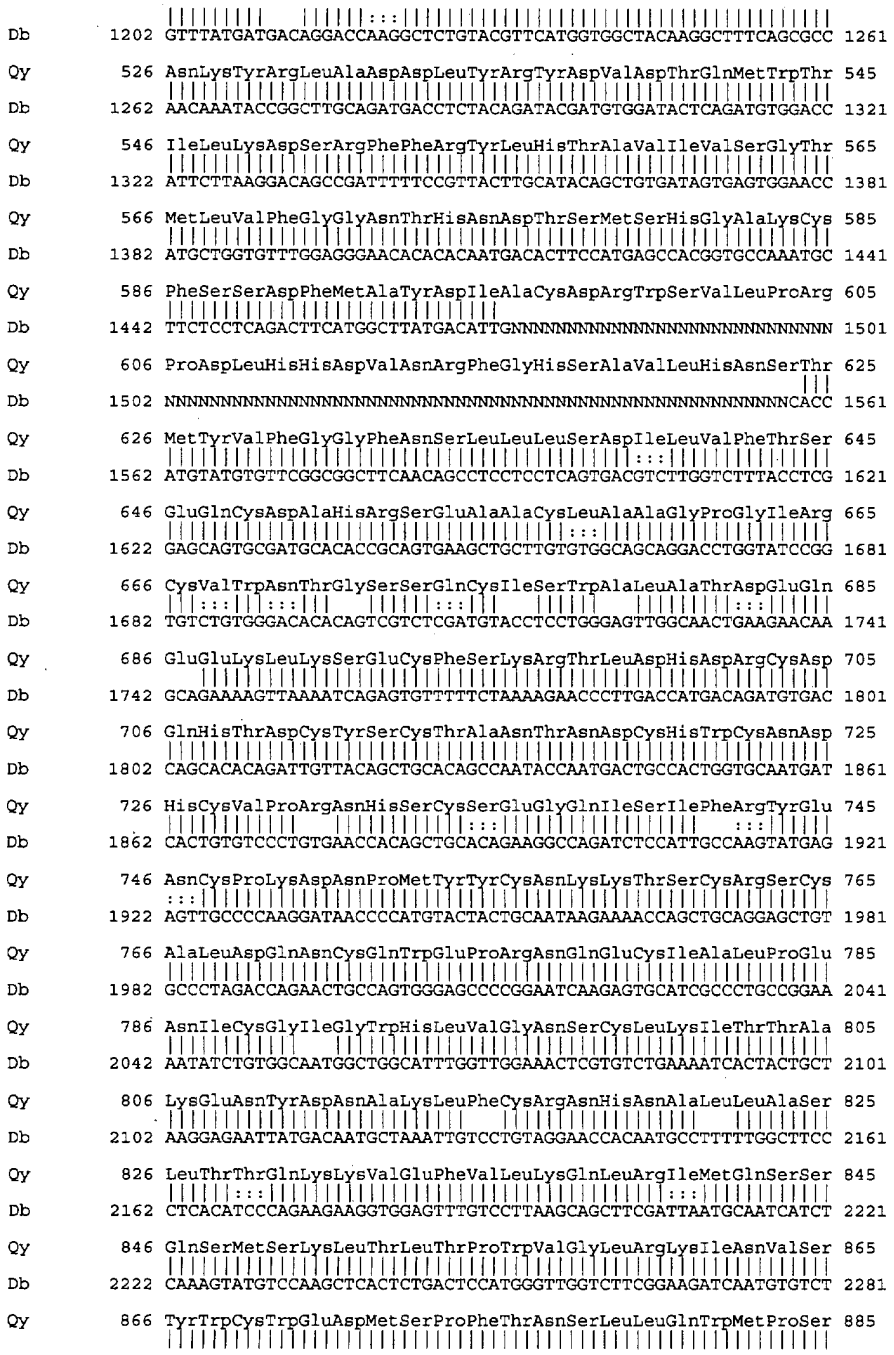
Qy        446 GlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValIlePhe 465

Db       1022 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1081

Qy        466 GlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsn 485
                                                                  ||||||
Db       1082 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNAAAAGAAC 1141

Qy        486 ThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisSerSer 505
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1142 ACATGGAGTATATTACATACTCAGGGTGCTCTTGTGCAAGGGGGTTATGGCCACAGTAGT 1201

Qy        506 ValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAla 525
```


Db 1202 GTTTATGATGACAGGACCAAGGCTCTGTACGTTCATGGTGGCTACAAGGCTTTCAGCGCC 1261
Qy 526 AsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetTrpThr 545
Db 1262 AACAAATACCGGCTTGCAGATGACCTCTACAGATACGATGTGGATACTCAGATGTGGACC 1321
Qy 546 IleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerGlyThr 565
Db 1322 ATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCATACAGCTGTGATAGTGAGTGGAACC 1381
Qy 566 MetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCys 585
Db 1382 ATGCTGGTGTTTGGAGGGAACACACACAATGACACTTCCATGAGCCACGGTGCCAAATGC 1441
Qy 586 PheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuProArg 605
Db 1442 TTCTCCTCAGACTTCATGGCTTATGACATTGNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1501
Qy 606 ProAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThr 625
Db 1502 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCACC 1561
Qy 626 MetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSer 645
Db 1562 ATGTATGTGTTCGGCGGCTTCAACAGCCTCCTCCTCAGTGACGTCTTGGTCTTTACCTCG 1621
Qy 646 GluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArg 665
Db 1622 GAGCAGTGCGATGCACACCGCAGTGAAGCTGCTTGTGTGGCAGCAGGACCTGGTATCCGG 1681
Qy 666 CysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGln 685
Db 1682 TGTCTGTGGGACACACAGTCGTCTCGATGTACCTCCTGGGAGTTGGCAACTGAAGAACAA 1741
Qy 686 GluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArgCysAsp 705
Db 1742 GCAGAAAAGTTAAAATCAGAGTGTTTTTCTAAAAGAACCCTTGACCATGACAGATGTGAC 1801
Qy 706 GlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAsp 725
Db 1802 CAGCACACAGATTGTTACAGCTGCACAGCCAATACCAATGACTGCCACTGGTGCAATGAT 1861
Qy 726 HisCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGlu 745
Db 1862 CACTGTGTCCCTGTGAACCACAGCTGCACAGAAGGCCAGATCTCCATTGCCAAGTATGAG 1921
Qy 746 AsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArgSerCys 765
Db 1922 AGTTGCCCCAAGGATAACCCCATGTACTACTGCAATAAGAAAACCAGCTGCAGGAGCTGT 1981
Qy 766 AlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuProGlu 785
Db 1982 GCCCTAGACCAGAACTGCCAGTGGGAGCCCCGGAATCAAGAGTGCATCGCCCTGCCGGAA 2041
Qy 786 AsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThrThrAla 805
Db 2042 AATATCTGTGGCAATGGCTGGCATTTGGTTGGAAACTCGTGTCTGAAAATCACTACTGCT 2101
Qy 806 LysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSer 825
Db 2102 AAGGAGAATTATGACAATGCTAAATTGTCCTGTAGGAACCACAATGCCTTTTTGGCTTCC 2161
Qy 826 LeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGlnSerSer 845
Db 2162 CTCACATCCCAGAAGAAGGTGGAGTTTGTCCTTAAGCAGCTTCGATTAATGCAATCATCT 2221
Qy 846 GlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSer 865
Db 2222 CAAAGTATGTCCAAGCTCACTCTGACTCCATGGGTTGGTCTTCGGAAGATCAATGTGTCT 2281
Qy 866 TyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSer 885
Db 2282 TACTGGTGCTGGGAGGATATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCT 2341
Qy 886 GluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLys 905
Db 2342 GAGCCCAGTGATGCTGGCTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGATTAAAG 2401
Qy 906 AlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSer 925
Db 2402 GCTGCAACCTGCATCAACCCTCTCAATGGCAGCGTCTGTGAAAGGCCTGCAAACCACAGT 2461
Qy 926 AlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGly 945
Db 2462 GCCAAGCAGTGCCGGACACCATGTGCCCTGCGGACAGCGTGTGGCGAGTGCACTAGCAGC 2521
Qy 946 SerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrVal 965
Db 2522 AGCTCGGAGTGCATGTGGTGCAGTAACATGAAGCAGTGTGTGGACTCCAATGCCTACGTG 2581
Qy 966 AlaSerProProPheGlyGlnCysMetGluTrpTyrThrMetSerThrCysProProGlu 985
Db 2582 GCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCAGCTGCCCACCTGAA 2641
Qy 986 AsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCys 1005
Db 2642 AATTGCTCTGGCTACTGTACCTGCAGCCATTGCTTGGAGCAGCCAGGCTGTGGTTGGTGT 2701
Qy 1006 ThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyProVal 1025
Db 2702 ACTGATCCTAGCAATACTGGGAAAGGAAAATGTATTGAGGGCAGCTATAAAGGACCTGTG 2761
Qy 1026 LysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerSer 1045
Db 2762 AAGATGCCGTCACAGGCCTCTGCAGGAAATGTGTATCCACAGCCCCTTCTGAACTCCAGC 2821
Qy 1046 MetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnCys 1065
Db 2822 ATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAGTGC 2881
Qy 1066 AsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrThr 1085
Db 2882 AACGGACACAGCAAATGCATCAACCAGAGTATCTGTGAGAAGTGTGAGGACCTGACCACG 2941
Qy 1086 GlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyGlyLys 1105
Db 2942 GGCAAGCACTGCGAGACCTGCATATCTGGCTTCTATGGTGACCCGACTAATGGAGGCAAA 3001
Qy 1106 CysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCys 1125
Db 3002 TGTCAGCCATGCAAGTGCAATGGGCACGCATCACTGTGCAACACCAACACCGGCAAGTGC 3061
Qy 1126 PheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGluAsnArg 1145
Db 3062 TTCTGTACCACCAAAGGTGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGA 3121
Qy 1146 TyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrGlnPhe 1165
Db 3122 TACCAGGGAAACCCTCTCAAAGGAACATGCTACTATACCCTTCTCATTGACTATCAGTTC 3181
Qy 1166 ThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIleAsnPheValAlaThr 1185
Db 3182 ACCTTTAGCCTGTCCCAGGAAGACGACCGCTACTACACAGCCATCAACTTTGTGGCTACT 3241
Qy 1186 ProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeu 1205
Db 3242 CCTGATGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAAAACTTCAACCTC 3301
Qy 1206 AsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProVal 1225
Db 3302 AACATCACCTGGGCCACCAGCTTCCCAGCCGGAACCCAGACTGGAGAAGAGGTGCCTGTT 3361
Qy 1226 ValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPheAspPhe 1245
Db 3362 GTTTCAAAAACCAACATCAAGGAATACAAAGATAGCTTCTCTAATGAGAAATTTGATTTT 3421

```
Qy       1246 ArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpProIleLys 1265
Db       3422 CGCAACCATCCAAACATCACTTTCTTTGTTTATGTCAGTAATTTCACTTGGCCCATCAAA 3481

Qy       1266 IleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePheValThr 1285
Db       3482 ATTCAGATTGCCTTCTCCCAGCACAGCAACTTCATGGACCTGGTACAGTTCTTCGTGACT 3541

Qy       1286 PhePheSerCysPheLeuSerLeuLeuLeuValAlaAlaValValTrpLysIleLysGln 1305
Db       3542 TTCTTCAGTTGTTTCCTCTCGCTGCTTCTGGTGGCTGCAGTGGTCTGGAAGATCAAGCAG 3601

Qy       1306 SerCysTrpAlaSerArgArgArgGluGlnLeuLeuArgGluMetGlnGlnMetAlaSer 1325
Db       3602 AGCTGTTGGGCATCCAGGCGGAGAGAGCAACTTCTTCGGGAGATGCAACAGATGGCCAGC 3661

Qy       1326 ArgProPheAlaSerValAsnValAlaLeuGluThrAspGluGluProProAspLeuIle 1345
Db       3662 CGCCCCTTTGCTTCTGTAAACGTTGCCTTGGAAACAGATGAAGAACCTCCTGATCTCATT 3721

Qy       1346 GlyGlySerIleLysThrValProLysProIleAlaLeuGluProCysPheGlyAsnLys 1365
Db       3722 GGGGGAAGTATAAAGACCGTTCCTAAGCCCATTGCCCTGGAGCCCTGCTTTGGTAACAAA 3781

Qy       1366 AlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyGlyIleProProPro 1385
Db       3782 GCCGCAGTCCTCTCTGTATTCGTGAGGCTCCCTCGAGGACTGGGAGGAATCCCTCCTCCT 3841

Qy       1386 GlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetProIle 1405
Db       3842 GGTCAGTCAGGTCTCGCTGTGGCCAGTGCCCTGGTGGACATTTCTCAGCAGATGCCAATA 3901

Qy       1406 ValTyrLysGluLysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnPro 1425
Db       3902 GTGTACAAGGAGAAGTCAGGAGCTGTAAGAAACCGGAAGCAGCAGCCGCCTGCACAGCCT 3961

Qy       1426 GlyThrCysIle 1429
Db       3962 GGAACCTGCATT 3973
```

```
RESULT 3
AY418588
LOCUS       AY418588               3671 bp    DNA     linear   GSS 17-DEC-2003
DEFINITION  Pan troglodytes ATRN gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY418588
VERSION     AY418588.1  GI:39774548
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 3671)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 3671)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     These sequences were made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .3671
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
     gene            <1. .>3671
                     /gene="ATRN"
                     /locus_tag="HCM6607"
ORIGIN
```

```
Alignment Scores:
Pred. No.:                0              Length:          3671
Score:                    5018.00        Matches:         950
Percent Similarity:       77.76%         Conservative:    1
Best Local Similarity:    77.68%         Mismatches:      272
Query Match:              63.40%         Indels:          0
DB:                       29             Gaps:            0

US-09-787-097-12 (1-1429) x AY418588 (1-3671)

Qy        166 ProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHis 185
Db          2 CCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCAT 61

Qy        186 LeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaAlaPheSerGlyLeu 205
Db         62 TTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGNNNNNNN 121

Qy        206 IleValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGlyTyr 225
Db        122 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 181

Qy        226 AlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThrTyr 245
Db        182 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 241

Qy        246 SerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSer 265
Db        242 NNTTTTGGTATGTGTCCGAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAGC 301

Qy        266 SerGluThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIlePro 285
Db        302 AGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCT 361

Qy        286 HisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspValArg 305
Db        362 CACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGA 421

Qy        306 GlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAlaAsn 325
Db        422 GGATGCTCCTGCTTCTCAGACTGGCAGGNNNCTGGATGTTCAGTTCCTGTACCAGCTAAC 481

Qy        326 GlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHisLys 345
Db        482 CAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAAA 541

Qy        346 AlaValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHisSerAsp 365
Db        542 GCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGAT 601

Qy        366 TyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSer 385
Db        602 TATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCT 661

Qy        386 ValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyr 405
Db        662 GTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATACAAGNNNNNNNNNNN 721

Qy        406 MetTyrGlyGlyLysIleAspProThrGlyAsnValThrAsnGluLeuArgValPheHis 425
Db        722 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 781

Qy        426 IleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaValVal 445
```

```
Db      782 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 841

Qy      446 GlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValIlePhe 465
Db      842 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 901

Qy      466 GlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsn 485
Db      902 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNATAAGAAC 961

Qy      486 ThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisSerSer 505
Db      962 ACGTGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGT 1021

Qy      506 ValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAla 525
Db     1022 GTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCC 1081

Qy      526 AsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetTrpThr 545
Db     1082 AATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACC 1141

Qy      546 IleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerGlyThr 565
Db     1142 ATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACC 1201

Qy      566 MetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCys 585
Db     1202 ATGCTGGTGTTTGGAGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATGC 1261

Qy      586 PheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuProArg 605
Db     1262 TTCTCTTCAGATTTCATGGCCTATGACATTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1321

Qy      606 ProAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThr 625
Db     1322 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCACC 1381

Qy      626 MetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSer 645
Db     1382 ATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCG 1441

Qy      646 GluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArg 665
Db     1442 GAACAGTGTGATGGGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATCCGG 1501

Qy      666 CysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGln 685
Db     1502 TGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAA 1561

Qy      686 GluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArgCysAsp 705
Db     1562 GAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGAC 1621

Qy      706 GlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAsp 725
Db     1622 CAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGCAATGAC 1681

Qy      726 HisCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGlu 745
Db     1682 CATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAG 1741

Qy      746 AsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArgSerCys 765
Db     1742 AATTGCCCCAAGGATAACCCCATGTNNNACTGTAACAAGAAGACCAGCTGCAGGAGCTGT 1801

Qy      766 AlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuProGlu 785
Db     1802 GCNNNGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAA 1861

Qy      786 AsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThrThrAla 805
Db     1862 AATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCC 1921

Qy      806 LysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSer 825
Db     1922 AAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCT 1981

Qy      826 LeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGlnSerSer 845
Db     1982 CTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGNNAATAATGCAGTCATCT 2041

Qy      846 GlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSer 865
Db     2042 CNNNNNNNNTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCC 2101

Qy      866 TyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSer 885
Db     2102 TACTGGTGCTGGGAAGATATGTCTCCATTTACAAATAGTTTACTACAGTGGATGCCGTCT 2161

Qy      886 GluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLys 905
Db     2162 GAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAG 2221

Qy      906 AlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSer 925
Db     2222 GCTGCAACCTGCATCAACCCACTCAATGGTAGCGTCTGTGAAAGGCCTGCAAACCACAGT 2281

Qy      926 AlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGly 945
Db     2282 GCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGGC 2341

Qy      946 SerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrVal 965
Db     2342 AGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTACGTG 2401

Qy      966 AlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSerThrCysProProGlu 985
Db     2402 GCCTCCTTCCCTTTCGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCNNNNN 2461

Qy      986 AsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCys 1005
Db     2462 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2521

Qy     1006 ThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyProVal 1025
Db     2522 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2581

Qy     1026 LysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerSer 1045
Db     2582 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2641

Qy     1046 MetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnCys 1065
Db     2642 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCTTGCCAATGC 2701

Qy     1066 AsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrThr 1085
Db     2702 AACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCACA 2761

Qy     1086 GlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyGlyLys 1105
Db     2762 GGCAAGCACTGCGAGACCTGCATATCTGGCTTCTATGGTGATCCCACCANNNNNNNNNNN 2821

Qy     1106 CysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCys 1125
Db     2822 NNNNNNNCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGC 2881

Qy     1126 PheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGluAsnArg 1145
Db     2882 TTCTGCACCACCAAGGGCGTCANNNNNNNNNNNNNNNNNNNATGTGAGGTAGAAAATCGA 2941

Qy     1146 TyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrGlnPhe 1165
Db     2942 TACCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNATACTCTTCTTATTGACTATCAGTTC 3001
```

```
Qy       1166 ThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIleAsnPheValAlaThr 1185
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3002 ACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATTAATTTGTGGCTACT 3061

Qy       1186 ProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeu 1205
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3062 CCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCTC 3121

Qy       1206 AsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProVal 1225
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3122 AACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTT 3181

Qy       1226 ValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPheAspPhe 1245
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3182 GTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTT 3241

Qy       1246 ArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpProIleLys 1265
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3242 CGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAAA 3301

Qy       1266 IleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePheValThr 1285
             |||||||||||||||||||      ||||||||||||||||||||||||||||||||||||
Db       3302 ATTCAGATTGCCTTCTCTCNNNACAGCAATTTTATGGACCTGGTACAGTTCTTCGTGACT 3361

Qy       1286 PhePheSerCysPheLeuSerLeuLeuLeuValAlaAlaValValTrpLysIleLysGln 1305
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3362 TTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGTGGTTTGGAAGATCAAACAA 3421

Qy       1306 SerCysTrpAlaSerArgArgArgGluGlnLeuLeuArgGluMetGlnGlnMetAlaSer 1325
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3422 AGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGAGATGCAACAGATGGCCAGC 3481

Qy       1326 ArgProPheAlaSerValAsnValAlaLeuGluThrAspGluGluProProAspLeuIle 1345
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3482 CGTCCCTTTGCCTCCGTAAATGTCGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTATT 3541

Qy       1346 GlyGlySerIleLysThrValProLysProIleAlaLeuGluProCysPheGlyAsnLys 1365
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3542 GGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGAGCCATGTTTTGGCAACAAA 3601

Qy       1366 AlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyGlyIleProProPro 1385
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3602 GCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCTGGGTGGCATCCCTCCTCCT 3661

Qy       1386 GlyGlnSer 1388
             |||||||||
Db       3662 GGGCAGTCA 3670
```

```
RESULT 4
BX440935
LOCUS       BX440935               1201 bp    mRNA    linear   EST 15-MAY-2003
DEFINITION  BX440935 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
            CS0DF012YL19 5-PRIME, mRNA sequence.
ACCESSION   BX440935
VERSION     BX440935.1  GI:30781889
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 10212.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DF012CF10QP1&cluster=10212.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DF012CF10QP1.
FEATURES             Location/Qualifiers
     source          1. .1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DF012YL19"
                     /tissue_type="FETAL BRAIN"
                     /dev_stage="fetal"
                     /clone_lib="Homo sapiens FETAL BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN
```

```
Alignment Scores:
Pred. No.:                7.93e-135      Length:          1201
Score:                    1746.50        Matches:         323
Percent Similarity:       90.28%         Conservative:    2
Best Local Similarity:    89.72%         Mismatches:      20
Query Match:              22.07%         Indels:          19
DB:                       13             Gaps:            2

US-09-787-097-12 (1-1429) x BX440935 (1-1201)

Qy        780 CysIleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCys 799
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 TGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGT 117

Qy        800 LeuLysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHis 819
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 TTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCAC 177

Qy        820 AsnAlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeu 839
              ||||||||||||||||||||||||||||||      ||||||||||||||||||||||||||
Db        178 AATGCCCTTTTGGCTTCTCTTACAACCCTGATGAAGGTAGAATTTGTCCTTAAGCAGCTG 237

Qy        840 ArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeu 859
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 CGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTT 297

Qy        860 ArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeu 879
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 CGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTA 357

Qy        880 LeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluPro 899
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 CTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCC 417

Qy        900 SerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGlu 919
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 AGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAA 477

Qy        920 ArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCys 939
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 AGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGT 537

Qy        940 GlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysVal 959
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 GGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTG 597

Qy        960 AspSerAsnAlaTyrValAlaSerProProGlyGlnCysMetGluTrpTyrThrMet 979
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 GACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATG 657

Qy        980 SerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGln 999
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        658 AGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAA 717
```

```
Qy       1000 ProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGly 1019
Db        718 CCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGT 777

Qy       1020 SerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGln 1039
Db        778 TCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAG 837

Qy       1040 ProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHis 1059
Db        838 CCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCAC 897

Qy       1060 CysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLys 1079
Db        898 TGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAG 957

Qy       1080 CysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAsp 1099
Db        958 TGTGAGAACYGAACAMA-GGMAG-CACTGCGAGAACTGCATATCTGGCTC-TACGGKGAT 1014

Qy       1100 ProThrAsnGlyGlyLysCys---------------GlnProCysLysCysAsnGlyHis 1114
Db       1015 CCACC-AATGGAGGGAATGTMRCMAKSARTGCAWGGGCACCSTCTYTTTCAAACAAMACG 1073

Qy       1115 AlaSerLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGly 1134
Db       1074 GGAATKCTT-----------------------------TKCRCACAAGGSGTMAGGGGA 1103
```

RESULT 5
BU506373
LOCUS BU506373 964 bp mRNA linear EST 12-SEP-2002
DEFINITION AGENCOURT_10015485 NIH_MGC_94 Mus musculus cDNA clone IMAGE:6494441 5', mRNA sequence.
ACCESSION BU506373
VERSION BU506373.1 GI:22812606
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 964)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14050 row: k column: 18
High quality sequence stop: 592.
FEATURES Location/Qualifiers
source 1. .964
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6494441"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_94"
/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 3.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.: 2.47e-124 Length: 964

Score:	1619.00	Matches:	289
Percent Similarity:	93.48%	Conservative:	12
Best Local Similarity:	89.75%	Mismatches:	19
Query Match:	20.45%	Indels:	2
DB:	13	Gaps:	0

US-09-787-097-12 (1-1429) x BU506373 (1-964)

```
Qy        666 CysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGln 685
Db          1 TGTCTGTGGGACACACAGTCGTCTCGATGTACCTCCTGGGAGTTGGCAACTGAAGAACAA 60

Qy        686 GluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArgCysAsp 705
Db         61 GCAGAAAAGTTAAAATCAGAGTGTTTTTCTAAAAGAACCCTTGACCATGACAGATGTGAC 120

Qy        706 GlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAsp 725
Db        121 CAGCACACAGATTGTTACAGCTGCACAGCCAATACCAATGACTGCCACTGGTGCAATGAT 180

Qy        726 HisCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGlu 745
Db        181 CACTGTGTCCCTGTGAACCACAGCTGCACAGAAGGCCAGATCTCCATTGCCAAGTATGAG 240

Qy        746 AsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArgSerCys 765
Db        241 AGTTGCCCCAAGGATAACCCCATGTACTACTGCAATAAGAAAACCAGCTGCAGGAGCTGT 300

Qy        766 AlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuProGlu 785
Db        301 GCCCTAGACCAGAACTGCCAGTGGGAGCCCCGGAATCAAGAGTGCATCGCCCTGCCGGAA 360

Qy        786 AsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThrThrAla 805
Db        361 AATATCTGTGGCAATGGCTGGCATTTGGTTGGAAACTCGTGTCTGAAAATCACTACTGCT 420

Qy        806 LysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSer 825
Db        421 AAGGAGAATTATGACAATGCTAAATTGTCCTGTAGGAACCACAATGCCTTTTTGGCTTCC 480

Qy        826 LeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGlnSerSer 845
Db        481 CTCACATCCCAGAAGAAGGTGGAGTTTGTCCTTAAGCAGCTTCGATTAATGCAATCATCT 540

Qy        846 GlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSer 865
Db        541 CAAAGTATGTCCAAGCTCACTCTGACTCCATGGGTTGGTCTTCGGAAGATCAATGTGTCT 600

Qy        866 TyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSer 885
Db        601 TACTGGTGCTGGGAGGATATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCT 660

Qy        886 GluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArg-GlyLeuLy 905
Db        661 GAGCCCAGTGATGCTGGCTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGGATTAAA 720

Qy        905 sAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSe 925
Db        721 GGCTGCAACCTGCATCAACCCTCTCAATGGCAGCGTCTGTGAAAAGGCTGCAAACCACAG 780

Qy        925 rAlaLysGlnCysArgThrProCysAlaLeuArgThrAla-CysGlyAspCysThrSerG 945
Db        781 TGCCAAGCAGTGCCGGACACCATGTGCCCTGCGGACANCCTTGTGGCCAGTGCACTAGCA 840

Qy        945 lySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrV 965
Db        841 CCAGCTCGGAGTGCATGTGGTGCAGTAACATGAAGCAATGTGGGGACTCCAATGCCTACG 900

Qy        965 alAlaSerProProPheGlyGlnCysMetGluTrpTyrThrMetSerThrCysProProG 985
Db        901 TGGCCTCCTTCCNTTTTGNCCAANGNATTGAAAGGTATACGATGAGCAGCTGTCCAACTG 960

Qy        985 lu 985
```

```
Db        961 AA 962

RESULT 6
CB520374
LOCUS       CB520374                840 bp    mRNA    linear   EST 09-JUL-2003
DEFINITION  UI-M-GI0-cei-1-18-0-UI.r1 NIH_BMAP_GI0 Mus musculus cDNA clone
            IMAGE:6840067 5', mRNA sequence.
ACCESSION   CB520374
VERSION     CB520374.1  GI:29353729
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 840)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jim Lin, University of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/mousefl.html
             This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .840
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:6840067"
                     /tissue_type="whole brain"
                     /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_BMAP_GI0"
                     /note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel.First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site .Double strand cDNA was
                     size selected according to mRNA size fraction, ligated
                     with EcoR I adaptor, digested with NotI and then cloned
                     directionally into pYX-Asc vector . The library tag
                     sequence located between the Not I site and the polyA tail
                     is AGCGAGACAG. This library was created for the University
                     Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
                     Developing Mouse Nervous System', supported by National
                     Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
                     program coordinator."
ORIGIN

Alignment Scores:
Pred. No.:              9.78e-114       Length:         840
Score:                  1491.00         Matches:        262
Percent Similarity:     96.07%          Conservative:   7
Best Local Similarity:  93.57%          Mismatches:     11
Query Match:            18.84%          Indels:         1
DB:                     14              Gaps:           0

US-09-787-097-12 (1-1429) x CB520374 (1-840)

Qy        718 AsnAspCysHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGluGly 737
              ||||||||||||||||||||||||||||||||||  ||||||||||:::||||||
Db          2 AATGACTGCCACTGGTGCAATGATCACTGTGTCCCTGTGAACCACAGCTGCACAGAAGGC 61

Qy        738 GlnIleSerIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsn 757
              ||||||||||||  :::||||||:::||||||||||||||||||||||||||||||||||
Db         62 CAGATCTCCATTGCCAAGTATGAGAGTTGCCCCAAGGATAACCCCATGTACTACTGCAAT 121

Qy        758 LysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsn 777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 AAGAAAACCAGCTGCAGGAGCTGTGCCCTAGACCAGAACTGCCAGTGGGAGCCCCGGAAT 181

Qy        778 GlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsn 797
              ||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||
Db        182 CAAGAGTGCATCGCCCTGCCGGAAAATATCTGTGGCAATGGCTGGCATTTGGTTGGAAAC 241

Qy        798 SerCysLeuLysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArg 817
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 TCGTGTCTGAAAATCACTACTGCTAAGGAGAATTATGACAATGCTAAATTGTCCTGTAGG 301

Qy        818 AsnHisAsnAlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLys 837
              ||||||||||||  ||||||||||||||:::||||||||||||||||||||||||||||||
Db        302 AACCACAATGCCTTTTTGGCTTCCCTCACATCCCAGAAGAAGGTGGAGTTTGTCCTTAAG 361

Qy        838 GlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpVal 857
              ||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 CAGCTTCGATTAATGCAATCATCTCAAAGTATGTCCAAGCTCACTCTGACTCCATGGGTT 421

Qy        858 GlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsn 877
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 GGTCTTCGGAAGATCAATGTGTCTTACTGGTGCTGGGAGGATATGTCTCCATTCACAAAT 481

Qy        878 SerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSer 897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 AGTTTGCTGCAGTGGATGCCATCTGAGCCCAGTGATGCTGGCTTCTGTGGGATCTTGTCA 541

Qy        898 GluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerVal 917
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 GAGCCTAGTACTCGGGGATTAAAGGCTGCAACCTGCATCAACCCTCTCAATGGCAGCGTC 601

Qy        918 CysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThr 937
              ||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db        602 TGTGAAAGGCCTGCANACCACAGTGCCAAGCAGTGCCGGACACCATGTGCCCTGCGGACA 661

Qy        938 AlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGln 957
              |||||||||:::||||||||  ||||||||||||||||||||  |||||||||
Db        662 GCGTGTGGCGAGTGCACTAGCAGCAGCTCGGAGTGCATGTGGTGCNAGTACATGAAGCAG 721

Qy        958 CysValAspSerAsnAlaTyrValAlaSerProProPheGlyGlnCysMetGluTrpTyr 977
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 TGTGTGGACTCCAATGCCTACGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 781

Qy        978 ThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeu 997
              |||||||||:::||||||||      ||||||||||||||||||||||||||||||||||
Db        782 ACGATGAGCAGCTGCCCACCTGA-NATTGCTCTGGCTACTGTACCTGCAGCCATTGCTTG 840

RESULT 7
BG678679
LOCUS       BG678679                933 bp    mRNA    linear   EST 01-MAY-2001
DEFINITION  602624477F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4749427 5',
            mRNA sequence.
ACCESSION   BG678679
VERSION     BG678679.1  GI:13910076
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 933)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: James Cleaver, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
```

```
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLAM10602  row: j  column: 20
           High quality sequence stop: 762.
FEATURES             Location/Qualifiers
     source          1. .933
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4749427"
                     /tissue_type="squamous cell carcinoma"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NCI_CGAP_Skn4"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.5kb.  Library  constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
ORIGIN

Alignment Scores:
Pred. No.:                7.56e-108      Length:          933
Score:                    1421.50        Matches:         258
Percent Similarity:       95.93%         Conservative:    1
Best Local Similarity:    95.56%         Mismatches:      10
Query Match:              17.96%         Indels:          2
DB:                       12             Gaps:            1

US-09-787-097-12 (1-1429) x BG678679 (1-933)

Qy       1000 ProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGly 1019
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 CCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGT 62

Qy       1020 SerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGln 1039
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 TCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAG 122

Qy       1040 ProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHis 1059
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 CCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCAC 182

Qy       1060 CysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLys 1079
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 TGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAG 242

Qy       1080 CysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAsp 1099
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 TGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGAT 302

Qy       1100 ProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsn 1119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 CCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAAC 362

Qy       1120 ThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeu 1139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 ACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTA 422

Qy       1140 CysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeu 1159
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 TGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTT 482

Qy       1160 LeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAla 1179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 CTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCT 542

Qy       1180 IleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAla 1199
              |||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||
Db        543 ATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTCGGACATGTTCATCAATGCC 602

Qy       1200 SerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAla 1219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 TCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCT 662

Qy       1220 GlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSer 1239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 GGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCT 722

Qy       1240 AsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsn 1259
              |||||||||  |||   |||   |||   ||||||||||||   ||||||||||||||||
Db        723 AATGAGAAGTCCGATTCTCGCAAGCACCA-AATATCACTTTCTGTGTGTATGTCAGTAAT 781

Qy       1260 ---PheThrTrpProIleLysIleGlnIle 1268
                  |||   |||||||||:::
Db        782 TGTCACTTGGGCCCATCGAAAATTCAGGTG 811

RESULT 8
AK050882
LOCUS       AK050882               1545 bp    mRNA    linear   HTC 20-SEP-2003
DEFINITION  Mus musculus 9 days embryo whole body cDNA, RIKEN full-length
            enriched library, clone:D030029G14 product:weakly similar to
            DJ741H3.1.2 (ATTRACTIN (WITH DIPEPTIDYLPEPTIDASE IV ACTIVITY)
            MEMBRANE BOUND ISOFORM (KIAA0548)) (FRAGMENT) [Homo sapiens], full
            insert sequence.
ACCESSION   AK050882
VERSION     AK050882.1  GI:26341559
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6  (bases 1 to 1545)
  AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
            Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
            Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
            Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
            Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
```

Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.

TITLE Direct Submission

JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.

FEATURES Location/Qualifiers

source 1. .1545
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:D030029G14"
/db_xref="MGI:2418771"
/db_xref="taxon:10090"
/clone="D030029G14"
/tissue_type="whole body"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="9 days embryo"

CDS <1. .1219
/note="unnamed protein product; putative
weakly similar to DJ741H3.1.2 (ATTRACTIN (WITH
DIPEPTIDYLPEPTIDASE IV ACTIVITY) MEMBRANE BOUND ISOFORM
(KIAA0548)) (FRAGMENT) [Homo sapiens] (SPTR|CAC32457,
evidence: FASTY, 56%ID, 35.3%length, match=1200)"
/codon_start=2
/protein_id="BAC34442.1"
/db_xref="GI:26341560"
/translation="CISASSNCSTSVRNYTKCHIRNEQICNKLTSCKSCSLNLNCQWD
QRQQECQALPAHLCGEGWNHVGDACLRINSSRESYDNAKLYCYNLSGNLASLTTSKEV
EFVLDEIQKFTQQKVSPWVGLRKINISYWGWEDMSPFTNTSLQWLPGEPNDSGFCAYL
ERAAVAGLKANPCTSMADGLVCEKPVVSPNQNARPCKKPCSLRTSCANCTSSGMECMW
CSSTKRCVDSNAYIISFPYGQCLEWQTATCSPQNCSGLRTCGQCLEQPGCGWCNDPSN
TGRGYCIEGSSRGPMKLAGVHNSDVVLDTSLCPKEKNYEWSFIQCPACQCNGHSTCIN
NNVCEQCKNLTTGRQCQECMPGYYGDPTNGGQCTACTCGGHANVCHLHTGKCFCTTKG
IKGDQCQLYCQFF"

polyA_signal 1525. .1530
/note="putative"

polyA_site 1545
/note="putative"

ORIGIN

Alignment Scores:

Pred. No.:	7.69e-106	Length:	1545
Score:	1402.00	Matches:	234
Percent Similarity:	73.75%	Conservative:	75
Best Local Similarity:	55.85%	Mismatches:	90
Query Match:	17.71%	Indels:	20
DB:	11	Gaps:	7

US-09-787-097-12 (1-1429) x AK050882 (1-1545)

```
Qy        727 CysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGluAsn 746
              |||:::      :::   :::||||||           |||:::      |||
Db          2 TGCATCTCAGCCAGCAGCAACTGCAGC---------ACGTCTGTCAGAAACTACACTAAA 52

Qy        747 CysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArgSerCysAla 766
              |||         |||            |||||||||   |||||||||:::||||||:::

Db         53 TGTCATATAAGAAATGAGCAG---ATTTGTAACAAACTTACAAGCTGTAAAAGCTGTTCA 109

Qy        767 LeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuProGluAsn 786
              |||:::   |||||||||||:::   |||   |||||||||   |||||||||   :::

Db        110 CTCAACTTGAATTGCCAGTGGGATCAGCGGCAGCAGGAATGTCAGGCTTTGCCAGCTCAC 169

Qy        787 IleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThrThrAlaLys 806
              :::||||||   ||||||:::   |||||:::::::||||||:::|||   :::::::::

Db        170 CTTTGTGGAGAAGGCTGGAATCATGTTGGGGACGCTTGTCTTCGAATCAATTCCAGTCGA 229

Qy        807 GluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSerLeu 826
              |||:::|||||||||||||||||:::|||   |||   :::       ||||||||||||

Db        230 GAAAGCTATGATAATGCCAAACTTTATTGCTATAACCTCAGTGGAAATCTCGCCTCCCTG 289

Qy        827 ThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGlnSerSerGln 846
              ||||||   |||:::|||||||||||||||   ::::::                  |||

Db        290 ACCACGTCCAAGGAGGTGGAGTTTGTGTTGGATGAAATA----------------CAG 331

Qy        847 SerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyr 866
                  ::::::      ::::::||||||||||||||||||||||||||:::||||||

Db        332 AAGTTCACACAGCAGAAAGTGTCACCGTGGGTAGGCCTACGCAAGATCAACATATCCTAC 391

Qy        867 TrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGlu 886
              |||   ||||||||||||||||||||||||||:::   |||||||||:::|||   |||

Db        392 TGGGGATGGGAGGACATGTCTCCTTTCACAAATACAAGCCTGCAGTGGCTTCCTGGTGAG 451

Qy        887 ProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAla 906
              |||:::|||:::|||||||||       |||          :::       ||||||||||||

Db        452 CCAAATGACTCTGGATTCTGTGCCTACTTAGAAAGGGCTGCAGTGGCAGGGTTAAAAGCA 511

Qy        907 AlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAla---------Asn 923
                  |||   :::      :::|||   |||||||||:::|||                |||

Db        512 AACCCTTGCACATCCATGGCAGATGGACTCGTTTGTGAAAAGCCTGTAGTAAGCCCAAAT 571

Qy        924 HisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThr 943
              :::|||:::   |||:::   |||||||:::|||||||||:::|||   :::|||||||

Db        572 CAGAACGCGAGGCCGTGCAAGAAGCCGTGCTCCCTGAGGACCTCCTGCGCCAACTGCACG 631

Qy        944 SerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAla 963
              |||         |||||||||||||||||:::   |||:::||||||||||||||||||

Db        632 AGCAGCGGCATGGAGTGCATGTGGTGCAGCAGCACGAAGCGCTGCGTGGACTCCAACGCT 691

Qy        964 TyrValAlaSerProProPheGlyGlnCysMetGluTrpTyrThrMetSerThrCysPro 983
              |||:::   |||||||||:::|||||||:::||||||   |||   :::||||||

Db        692 TACATCATCTCCTTTCCCTACGGACAGTGCCTGGAGTGGCAGACT---GCCACCTGCTCG 748

Qy        984 ProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGly 1003
              |||:::||||||||||||       ||||||       ||||||||||||||||||||||

Db        749 CCTCAAAATTGTTCTGGGTTAAGAACCTGTGGACAGTGCTTGGAGCAGCCTGGGTGTGGC 808

Qy       1004 TrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGly 1023
              ||||||   |||||||||||||||||:::|||   |||||||||||||   :::|||

Db        809 TGGTGCAACGATCCTAGTAACACAGGAAGAGGCTATTGCATCGAAGGGTCTTCCCGGGGC 868

Qy       1024 ProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsn 1043
              |||:::|||:::                |||                :::|||:::

Db        869 CCAATGAAACTC--------------GCGGGGGTCCACAACAGTGACGTGGTTCTTGAC 913

Qy       1044 SerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCys 1063
              :::|||:::|||  ::::::      |||   ||||||||||||   ||||||||||||

Db        914 ACCAGCCTCTGCCCCAAGGAGAAGAACTACGAGTGGTCTTTTATCCAGTGTCCAGCTTGC 973

Qy       1064 GlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeu 1083
              ||||||||||||||||   |||||||||   ::::::||||||:::|||:::||||||

Db        974 CAGTGTAATGGACACAGCACGTGCATCAACAACAACGTCTGTGAGCAGTGTAAGAATCTC 1033

Qy       1084 ThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGly 1103
              |||||||||:::   |||:::   |||:::   |||:::||||||||||||||||||||

Db       1034 ACCACTGGGCGACAGTGTCAGGAATGCATGCCAGGGTACTATGGAGACCCAACCAACGGT 1093

Qy       1104 GlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGly 1123
              |||:::|||      |||   |||   |||||||||:::::|||:::   :::|||||||

Db       1094 GGGCAGTGCACAGCTTGCACGTGCGGCGGCCATGCGAACGTCTGTCACCTGCACACGGGA 1153
```

```
Qy       1124 LysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeu---CysGlu 1141
              ||||||||||||||||||||||||:::|||||||||:::|||||||||   |||:::
Db       1154 AAGTGTTTTTGCACAACCAAGGGGATCAAGGGTGACCAGTGCCAGCTGTATTGTCAG 1210

RESULT 9
BY761276
LOCUS       BY761276                770 bp    mRNA    linear   EST 17-DEC-2002
DEFINITION  BY761276 RIKEN full-length enriched, bone marrow macrophage Mus
            musculus cDNA clone G530008F22 3', mRNA sequence.
ACCESSION   BY761276
VERSION     BY761276.1  GI:27196825
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 770)
  AUTHORS   Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
            Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
            Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
            Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
            Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
            Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
            Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
            Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
            Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
            Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
            Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
            Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
            Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
            Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
            Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
            Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
            Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
            Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
            Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
            Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
            Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
            Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
            Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
            Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
            Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
            Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
            Rogers,J., Birney,E. and Hayashizaki,Y.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
  MEDLINE   22354683
   PUBMED   12466851
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P.,
            Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F.,
            Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y.,
            Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M.,
            Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N.,
            Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M.,
            Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.
            Direct Submission
             Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. 10 (10), 1617-1630 (2000)
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
             cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
             Tissues were provided by David A. Hume ( Depts. of Biochemistry
            and Microbiology/Parasitology Institute for Molecular Bioscience
            University of Queensland Brisbane,Q 4072 Australia ) whose
            assistance we gratefully acknowledge.
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
FEATURES             Location/Qualifiers
     source          1. .770
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="G530008F22"
                     /tissue_type="bone marrow"
                     /cell_type="macrophage"
                     /clone_lib="RIKEN full-length enriched, bone marrow
                     macrophage"
ORIGIN

Alignment Scores:
Pred. No.:                6.07e-106      Length:          770
Score:                    1397.00        Matches:         240
Percent Similarity:       96.09%         Conservative:    6
Best Local Similarity:    93.75%         Mismatches:      10
Query Match:              17.65%         Indels:          0
DB:                       13             Gaps:            0

US-09-787-097-12 (1-1429) x BY761276 (1-770)

Qy        837 LysGlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrp 856
              ||||||:::   :::|||   |||||||||   ||||||:::|||||||||||   |||
Db          3 AAGCAGGTTGGATTGATGCNATCATCTCAATGTATGTCGCAGCTCACTCTGACTTCATGG 62

Qy        857 ValGlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThr 876
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 GTTGGTCTTCGGAAGATCAATGTGTCTTACTGGTGCTGGGAGGATATGTCTCCATTCACA 122

Qy        877 AsnSerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeu 896
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 AATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGTGATGCTGGCTTCTGTGGGATCTGN 182

Qy        897 SerGluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySer 916
              ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 TCAGAGCNTAGTACTCGGGGATTAAAGGCTGCAACCTGCATCAACCCTCTCAATGGCAGC 242

Qy        917 ValCysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArg 936
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCTGTGAAAGGCCTGCAAACCACAGTGCCAAGCAGTGCCGGACACCATGTGCCCTGCGG 302

Qy        937 ThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLys 956
              ||||||||||:::||||||||   ||||||||||||||||||||||||||||||||||||
Db        303 ACAGCGTGTGGCGAGTGCACTAGCAGCAGCTCGGAGTGCATGTGGTGCAGTAACATGAAG 362

Qy        957 GlnCysValAspSerAsnAlaTyrValAlaSerProProPheGlyGlnCysMetGluTrp 976
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 CAGTGTGTGGACTCCAATGCCTACGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGG 422

Qy        977 TyrThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCys 996
              ||||||||||||:::||||||||||||||||||||||||||||||||||||||||||||
Db        423 TATACGATGAGCAGCTGCCCACCTGAAAATTGCTCTGGCTACTGTACCTGCAGCCATTGC 482

Qy        997 LeuGluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCys 1016
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        483 TTGGAGCAGCCAGGCTGTGGTTGGTGTACTGATCCTAGCAATACTGGGAAAGGAAAATGT 542

Qy       1017 IleGluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPhe 1036
              ||||||||||||||||||||||||||||||||||||||||||||||||||      ||||||
Db        543 ATTGAGGGCAGCTATAAAGGACCTGTGAAGATGCCGTCACAGGCCTCTGCAGGAAATGTG 602

Qy       1037 TyrProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSer 1056
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 TATCCACAGCCCCTTCTGAACTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCT 662

Qy       1057 PheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIle 1076
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 TTCATTCACTGTCCAGCTTGCCAGTGCAACGGACACAGCAAATGCATCAACCAGAGTATC 722

Qy       1077 CysGluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCys 1092
              ||||||||||||||:::||||||||||||||||||||||||||||||
Db        723 TGTGAGAAGTGTGAGGACCTGACCACGGGCAAGCACTGCGAGACCTGC 770
```

RESULT 10
BY761390

```
LOCUS       BY761390                735 bp    mRNA    linear   EST 17-DEC-2002
DEFINITION  BY761390 RIKEN full-length enriched, bone marrow macrophage Mus
            musculus cDNA clone G530012E16 3', mRNA sequence.
ACCESSION   BY761390
VERSION     BY761390.1  GI:27197002
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 735)
  AUTHORS   Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
            Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
            Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
            Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
            Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
            Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
            Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
            Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
            Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
            Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
            Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
            Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
            Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
            Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
            Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
            Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
            Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
            Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
            Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
            Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
            Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
            Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
            Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
            Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
            Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
            Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
            Rogers,J., Birney,E. and Hayashizaki,Y.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
  MEDLINE   22354683
   PUBMED   12466851
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P.,
            Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F.,
            Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y.,
            Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M.,
            Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N.,
            Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M.,
            Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.
            Direct Submission
             Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. 10 (10), 1617-1630 (2000)
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
             cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
             Tissues were provided by David A. Hume ( Depts. of Biochemistry
            and Microbiology/Parasitology Institute for Molecular Bioscience
            University of Queensland Brisbane,Q 4072 Australia ) whose
            assistance we gratefully acknowledge.
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
FEATURES             Location/Qualifiers
     source          1. .735
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="G530012E16"
                     /tissue_type="bone marrow"
                     /cell_type="macrophage"
                     /clone_lib="RIKEN full-length enriched, bone marrow
                     macrophage"
ORIGIN
```

Alignment Scores:

Pred. No.:	4.76e-102	Length:	735
Score:	1350.00	Matches:	234
Percent Similarity:	96.73%	Conservative:	3
Best Local Similarity:	95.51%	Mismatches:	8
Query Match:	17.06%	Indels:	1
DB:	13	Gaps:	0

US-09-787-097-12 (1-1429) x BY761390 (1-735)

```
Qy        848 MetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrp 867
              |||||||||||||||||||||||||||||   ||||||||||||||||||||||||   |||
Db          2 ATGTCAAAGCTCACTCTGACTCCATGGGT-GGTCTTCGGAAGATCAATGTGTCTANCTGG 60

Qy        868 CysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluPro 887
              ||||||    ||||||||||||||||||||||    ||||||||||||||||||||||||
Db         61 TGCTGGNAGGATATGTCTCCATTCACAAATAGTNTGCTGCAGTGGATGCCATCTGAGCCC 120

Qy        888 SerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAla 907
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTGATGCTGGCTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGATTAAAGGCTGCA 180

Qy        908 ThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSerAlaLys 927
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCTGCATCAACCCTCTCAATGGCAGCGTCTGTGAAAGGCCTGCAAACCACAGTGCCAAG 240

Qy        928 GlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGlySerSer 947
              ||||||||||||||||||||||||||||||||||||:::|||||||||   ||||||
Db        241 CAGTGCCGGACACCATGTGCCCTGCGGACAGCGTGTGGCGAGTGCACTAGCAGCAGCTCG 300

Qy        948 GluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrValAlaSer 967
```

```
Db        301 GAGTGCATGTGGTGCAGTAACATGAAGCAGTGTGTGGACTCCAATGCCTACGTGGCCTCC 360

Qy        968 PheProPheGlyGlnCysMetGluTrpTyrThrMetSerThrCysProProGluAsnCys 987
Db        361 TTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCAGCTGCCCACCTGAAAATTGC 420

Qy        988 SerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCysThrAsp 1007
Db        421 TCTGGCTACTGTACCTGCAGCCATTGCTTGGAGCAGCCAGGCTGTGGTTGGTGTACTGAT 480

Qy       1008 ProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyProValLysMet 1027
Db        481 CCTAGCAATACTGGGAAAGGAAAATGTATTGAGGGCAGCTATAAAGGACCTGTGAAGATG 540

Qy       1028 ProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerSerMetCys 1047
Db        541 CCGTCACAGGCCTCTGCAGGAAATGTGTATCCACAGCCCCTTCTGAACTCCAGCATGTGT 600

Qy       1048 LeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnCysAsnGly 1067
Db        601 CTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAGTGCAACGGA 660

Qy       1068 HisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrThrGlyLys 1087
Db        661 CACAGCAAATGCATCAACCAGAGTATCTGTGAGAAGTGTGAGGACCTGACCACGGGCAAG 720

Qy       1088 HisCysGluThrCys 1092
Db        721 CACTGCGAGACCTGC 735

RESULT 11
CF745736
LOCUS       CF745736                777 bp    mRNA    linear   EST 10-OCT-2003
DEFINITION  UI-M-GV0-clw-d-20-0-UI.r1 NIH_BMAP_GV0 Mus musculus cDNA clone
            IMAGE:30620155 5', mRNA sequence.
ACCESSION   CF745736
VERSION     CF745736.1  GI:37642076
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 777)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James Lin, University of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/mousefl.html
             This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1..777
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:30620155"
                     /tissue_type="whole brain"
                     /dev_stage="1,5, and 15 days newborn"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_BMAP_GV0"
                     /note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel.First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site. Double strand cDNA was
                     size selected according to mRNA size fraction,ligated with
                     EcoR I adaptor, digested with NotI and then cloned
                     directionally into pYX-Asc vector. The library tag
                     sequence located between the Not I site and the polyA tail
                     is CGAACTGAAT. This library was created for the University
                     Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
                     Developing Mouse Nervous System', supported by National
                     Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
                     program coordinator."
ORIGIN

Alignment Scores:
Pred. No.:              4.41e-99      Length:        777
Score:                  1315.00       Matches:       229
Percent Similarity:     93.80%        Conservative:  13
Best Local Similarity:  88.76%        Mismatches:    16
Query Match:            16.61%        Indels:        0
DB:                     14            Gaps:          0

US-09-787-097-12 (1-1429) x CF745736 (1-777)

Qy        566 MetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCys 585
Db          3 ATGCTGGTGTTTGGAGGGAACACACACAATGACACTTCCATGAGCCACGGTGCCAAATGC 62

Qy        586 PheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuProArg 605
Db         63 TTCTCCTCAGACTTCATGGCTTATGACATTGCTTGTGACCGATGGTCAGTGCTTCCCAGA 122

Qy        606 ProAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThr 625
Db        123 CCTGAGCTCCATCATGATGTCAACAGATTTGGCCATTCAGCAGTCTTGTACAACAGCACC 182

Qy        626 MetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSer 645
Db        183 ATGTATGTGTTCGGCGGCTTCAACAGCCTCCTCCTCAGTGACGTCTTGGTCTTTACCTCG 242

Qy        646 GluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArg 665
Db        243 GAGCAGTGCGATGCACACCGCAGTGAAGCTGCTTGTGTGGCAGCAGGACCTGGTATCCGG 302

Qy        666 CysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGln 685
Db        303 TGTCTGTGGGACACACAGTCGTCTCGATGTACCTCCTGGGAGTTGGCAACTGAAGAACAA 362

Qy        686 GluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArgCysAsp 705
Db        363 GCAGAAAAGTTAAAATCAGAGTGTTTTTCTAAAAGAACCCTTGACCATGACAGATGTGAC 422

Qy        706 GlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAsp 725
Db        423 CAGCACACAGATTGTTACAGCTGCACAGCCAATACCAATGACTGCCACTGGTGCAATGAT 482

Qy        726 HisCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGlu 745
Db        483 CACTGTGTCCCTGTGAACCACAGCTGCACAGAAGGCCAGATCTCCATTGCCAAGTATGAG 542

Qy        746 AsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArgSerCys 765
Db        543 AGTTGCCCCAAGGATAACCCCATGTACTACTGCAATAAGAAAACCAGCTGCAGGAGCTGT 602

Qy        766 AlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuProGlu 785
Db        603 GCCCTAGACCAGAACTGCCAGTGGGAGCCCCGGAATCAAGAGTGCATCGCCCTGCCGGAA 662

Qy        786 AsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThrThrAla 805
Db        663 AATATCTGTGGCAATGGCTGGCATTTNGGTTGGAACTCGTGTCTGAAAATCACTACTGCT 722

Qy        806 LysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeu 823
```

```
Db        723 AAGGGAGATTATGACAATGCTAAATTGTCCTGTANNGANNCACATGCCTTTTTG 776

RESULT 12
BM783739
LOCUS       BM783739                681 bp    mRNA    linear   EST 05-MAR-2002
DEFINITION  K-EST0061730 S5SNU484 Homo sapiens cDNA clone S5SNU484-34-H05 5',
            mRNA sequence.
ACCESSION   BM783739
VERSION     BM783739.1  GI:19131971
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 681)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 34  row: H  column: 05
            High quality sequence stop: 681.
FEATURES             Location/Qualifiers
     source          1. .681
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="S5SNU484-34-H05"
                     /sex="M"
                     /tissue_type="Stomach"
                     /cell_type="Epithelial"
                     /cell_line="SNU-484"
                     /lab_host="Top10F'"
                     /clone_lib="S5SNU484"
                     /note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI;
                     Site_2: NotI; The poly (A)+ RNA was decapped with tabacco
                     acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
                     including EcoRI site by treatment of T4 RNA ligase. The
                     first strand cDNA was synthesized from oligo dT-selected
                     mRNA by priming with dT-tailed vector. The dT-tailed
                     vector was adjusted to have about 60nt. The cDNA vector
                     was circularized with E. coli DNA ligase after digestion
                     of EcoRI which site is also included in vector. An RNA
                     strand converted to a DNA strand by Okayama-Berg method.
                     The obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation
                     method."
ORIGIN

Alignment Scores:
Pred. No.:                  9.21e-99        Length:         681
Score:                      1310.00         Matches:        227
Percent Similarity:         100.00%         Conservative:   0
Best Local Similarity:      100.00%         Mismatches:     0
Query Match:                16.55%          Indels:         0
DB:                         12              Gaps:           0

US-09-787-097-12 (1-1429) x BM783739 (1-681)

Qy        838 GlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpVal 857
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTC 60

Qy        858 GlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsn 877
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAAT 120

Qy        878 SerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSer 897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCA 180

Qy        898 GluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerVal 917
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTC 240

Qy        918 CysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThr 937
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACA 300

Qy        938 AlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGln 957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 360

Qy        958 CysValAspSerAsnAlaTyrValAlaSerProPheGlyGlnCysMetGluTrpTyr 977
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 420

Qy        978 ThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeu 997
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTG 480

Qy        998 GluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIle 1017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATA 540

Qy       1018 GluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyr 1037
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTAT 600

Qy       1038 ProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPhe 1057
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 660

Qy       1058 IleHisCysProAlaCysGln 1064
              |||||||||||||||||||||
Db        661 ATTCACTGTCCAGCTTGCCAA 681

RESULT 13
BM783258
LOCUS       BM783258                680 bp    mRNA    linear   EST 05-MAR-2002
DEFINITION  K-EST0061157 S5SNU484 Homo sapiens cDNA clone S5SNU484-31-H08 5',
            mRNA sequence.
ACCESSION   BM783258
VERSION     BM783258.1  GI:19131490
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 680)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 31  row: H  column: 08
            High quality sequence stop: 680.
FEATURES             Location/Qualifiers
     source          1. .680
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
```

/clone="S5SNU484-31-H08"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/cell_line="SNU-484"
/lab_host="Top10F'"
/clone_lib="S5SNU484"
/note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI; Site_2: NotI; The poly (A)+ RNA was decapped with tabacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method."

ORIGIN

Alignment Scores:

Pred. No.:	2.41e-98	Length:	680
Score:	1305.00	Matches:	226
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	16.49%	Indels:	0
DB:	12	Gaps:	0

US-09-787-097-12 (1-1429) x BM783258 (1-680)

```
Qy        838 GlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpVal 857
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTC 60

Qy        858 GlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsn 877
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAAT 120

Qy        878 SerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSer 897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCA 180

Qy        898 GluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerVal 917
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTC 240

Qy        918 CysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThr 937
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACA 300

Qy        938 AlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGln 957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 360

Qy        958 CysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyr 977
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 420

Qy        978 ThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeu 997
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTG 480

Qy        998 GluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIle 1017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATA 540

Qy       1018 GluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyr 1037
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTAT 600

Qy       1038 ProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPhe 1057
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 660

Qy       1058 IleHisCysProAlaCys 1063
              ||||||||||||||||||
Db        661 ATTCACTGTCCAGCTTGC 678
```

RESULT 14
BM783773

```
LOCUS       BM783773                680 bp    mRNA    linear   EST 05-MAR-2002
DEFINITION  K-EST0061775 S5SNU484 Homo sapiens cDNA clone S5SNU484-36-D02 5',
            mRNA sequence.
ACCESSION   BM783773
VERSION     BM783773.1  GI:19132005
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 680)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 36  row: D  column: 02
            High quality sequence stop: 680.
FEATURES             Location/Qualifiers
     source          1. .680
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="S5SNU484-36-D02"
                     /sex="M"
                     /tissue_type="Stomach"
                     /cell_type="Epithelial"
                     /cell_line="SNU-484"
                     /lab_host="Top10F'"
                     /clone_lib="S5SNU484"
                     /note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI;
                     Site_2: NotI; The poly (A)+ RNA was decapped with tabacco
                     acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
                     including EcoRI site by treatment of T4 RNA ligase. The
                     first strand cDNA was synthesized from oligo dT-selected
                     mRNA by priming with dT-tailed vector. The dT-tailed
                     vector was adjusted to have about 60nt. The cDNA vector
                     was circularized with E. coli DNA ligase after digestion
                     of EcoRI which site is also included in vector. An RNA
                     strand converted to a DNA strand by Okayama-Berg method.
                     The obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation
                     method."
ORIGIN
```

Alignment Scores:

Pred. No.:	2.41e-98	Length:	680
Score:	1305.00	Matches:	226
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	16.49%	Indels:	0
DB:	12	Gaps:	0

US-09-787-097-12 (1-1429) x BM783773 (1-680)

```
Qy        838 GlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpVal 857
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db          1 CAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTC 60

Qy        858 GlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsn 877
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAAT 120

Qy        878 SerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSer 897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCA 180

Qy        898 GluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerVal 917
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTC 240

Qy        918 CysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThr 937
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACA 300

Qy        938 AlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGln 957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 360

Qy        958 CysValAspSerAsnAlaTyrValAlaSerProProGlyGlnCysMetGluTrpTyr 977
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 420

Qy        978 ThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeu 997
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACGATGAGCACCTGCCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTG 480

Qy        998 GluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIle 1017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATA 540

Qy       1018 GluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyr 1037
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTAT 600

Qy       1038 ProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPhe 1057
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 660

Qy       1058 IleHisCysProAlaCys 1063
              |||||||||||||||||
Db        661 ATTCACTGTCCAGCTTGC 678
```

RESULT 15
BM783788

```
LOCUS       BM783788                683 bp    mRNA    linear   EST 05-MAR-2002
DEFINITION  K-EST0061794 S5SNU484 Homo sapiens cDNA clone S5SNU484-36-E09 5',
            mRNA sequence.
ACCESSION   BM783788
VERSION     BM783788.1  GI:19132020
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 683)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 36  row: E  column: 09
            High quality sequence stop: 683.
```

```
FEATURES             Location/Qualifiers
     source          1. .683
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="S5SNU484-36-E09"
                     /sex="M"
                     /tissue_type="Stomach"
                     /cell_type="Epithelial"
                     /cell_line="SNU-484"
                     /lab_host="Top10F'"
                     /clone_lib="S5SNU484"
                     /note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI;
                     Site_2: NotI; The poly (A)+ RNA was decapped with tabacco
                     acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
                     including EcoRI site by treatment of T4 RNA ligase. The
                     first strand cDNA was synthesized from oligo dT-selected
                     mRNA by priming with dT-tailed vector. The dT-tailed
                     vector was adjusted to have about 60nt. The cDNA vector
                     was circularized with E. coli DNA ligase after digestion
                     of EcoRI which site is also included in vector. An RNA
                     strand converted to a DNA strand by Okayama-Berg method.
                     The obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation
                     method."
ORIGIN
```

```
Alignment Scores:
Pred. No.:                5.23e-98        Length:          683
Score:                    1301.00         Matches:         225
Percent Similarity:       99.56%          Conservative:    1
Best Local Similarity:    99.12%          Mismatches:      1
Query Match:              16.44%          Indels:          0
DB:                       12              Gaps:            0

US-09-787-097-12 (1-1429) x BM783788 (1-683)

Qy        838 GlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpVal 857
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTC 60

Qy        858 GlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsn 877
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAAT 120

Qy        878 SerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSer 897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCA 180

Qy        898 GluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerVal 917
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTC 240

Qy        918 CysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThr 937
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACA 300

Qy        938 AlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGln 957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 360

Qy        958 CysValAspSerAsnAlaTyrValAlaSerProProGlyGlnCysMetGluTrpTyr 977
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 420

Qy        978 ThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeu 997
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACGATGAGCACCTGCCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTG 480

Qy        998 GluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIle 1017
              ||||||||||||||||||||||||||||||||||||   :::||||||||||||||||||
Db        481 GAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGNCATACTGGCAAAGGGAAATGCATA 540
```

```
Qy       1018 GluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyr 1037
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTAT 600

Qy       1038 ProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPhe 1057
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 660

Qy       1058 IleHisCysProAlaCysGln 1064
              |||||||||||||||||||||
Db        661 ATTCACTGTCCAGCTTGCCAA 681
```

Search completed: March 1, 2004, 07:06:40
Job time : 6974 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 29, 2004, 11:51:51 ; Search time 1076 Seconds
(without alignments)
5641.891 Million cell updates/sec

Title: US-09-787-097-12
Perfect score: 7915
Sequence: 1 MVAAAAATEARLRRRTAATA.....KSGAVRNRKQQPPAQPGTCI 1429

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09787097/runat_26022004_062934_23410/app_query.fasta_1.1607
-DB=N_Geneseq_29Jan04 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09787097_@CGN_1_1_1137_@runat_26022004_062934_23410 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	7915	100.0	4290	3	AAZ52276	Aaz52276 Human mem
	2	7915	100.0	8561	5	AAS72659	Aas72659 DNA encod
	3	7696.5	97.2	8590	3	AAZ91920	Aaz91920 Human mah
	4	7478	94.5	4068	3	AAZ52275	Aaz52275 Human mem
	5	7342.5	92.8	8739	9	ADB53249	Adb53249 Primary r
	6	7294	92.2	8827	3	AAZ91916	Aaz91916 Wild type
	7	7089	89.6	3819	3	AAZ52277	Aaz52277 Human sol
	8	6870.5	86.8	4072	3	AAZ91921	Aaz91921 Human mah
	9	6612	83.5	3597	3	AAZ52274	Aaz52274 Human sol
	10	6612	83.5	3597	5	AAS72657	Aas72657 DNA encod
	11	6612	83.5	3597	9	ADD70997	Add70997 Human att
	12	5622	71.0	3490	5	AAS72660	Aas72660 DNA encod
	13	4637.5	58.6	4140	7	ABZ58709	Abz58709 Human hom
	14	4623	58.4	4134	7	ABZ58710	Abz58710 Murine ho
	15	4136.5	52.3	2625	3	AAZ91922	Aaz91922 Human mah
	16	3880	49.0	3609	6	ABS63434	Abs63434 Human cDN
	17	3452	43.6	6373	3	AAZ91919	Aaz91919 Human mah
	18	3435	43.4	2419	3	AAZ91917	Aaz91917 Murine ma
	19	2506.5	31.7	6517	4	AAH98748	Aah98748 Human EST
	20	2413.5	30.5	6532	5	AAS86766	Aas86766 DNA encod
	21	2232.5	28.2	4164	4	ABL09747	Abl09747 Drosophil
	22	2084	26.3	3165	5	AAS78781	Aas78781 DNA encod
	23	2084	26.3	3313	5	AAS76294	Aas76294 DNA encod
c	24	2032.5	25.7	10315	4	ABL09746	Abl09746 Drosophil
	25	1825	23.1	1012	2	AAX84568	Aax84568 Biorhythm
	26	1298	16.4	1654	7	ABX34730	Abx34730 Human mdd
	27	1259.5	15.9	1383	6	AAL54064	Aal54064 cDNA of h
	28	1226.5	15.5	1051	3	AAZ91918	Aaz91918 Murine ma
	29	977.5	12.3	8875	9	ADC51661	Adc51661 DNA encod
	30	977.5	12.3	9921	9	ADC51659	Adc51659 DNA encod
	31	950	12.0	6056	6	ABS51293	Abs51293 Human REM
	32	937	11.8	8865	4	ABL12959	Abl12959 Drosophil
	33	923	11.7	7244	7	ABX70916	Abx70916 Novel hum
	34	895	11.3	11502	4	ABL12958	Abl12958 Drosophil
	35	840	10.6	207433	5	ABZ72040	Abz72040 Gene 216
	36	840	10.6	207433	7	ABX74891	Abx74891 BAC1098L2
	37	766	9.7	90050	3	AAZ91925	Aaz91925 Wild type
	38	722	9.1	671	5	AAS72658	Aas72658 DNA encod
	39	708	8.9	982	6	ABL39784	Abl39784 Human NS
	40	684.5	8.6	657	7	ABZ58714	Abz58714 Upregulat
	41	557	7.0	668	4	AAS02409	Aas02409 Human sec
	42	553	7.0	292	3	AAA45505	Aaa45505 Human sec
c	43	536	6.8	481	4	AAL14686	Aal14686 Human bre
c	44	523	6.6	482	4	AAL23548	Aal23548 Human bre
	45	477	6.0	845	6	ABQ52394	Abq52394 Oligonucl

ALIGNMENTS

RESULT 1
AAZ52276
ID AAZ52276 standard; cDNA; 4290 BP.
XX
AC AAZ52276;
XX
DT 18-JUL-2000 (first entry)
XX
DE Human membrane attractin-2 cDNA.
XX
KW Human; membrane attractin-2; immune response; macrophage; monocyte;
KW T cell; immunostimulant; immunosuppressed patient; cancer;
KW immunodeficiency syndrome; transplant; autoimmune disease; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1. .4290
FT /*tag= a
FT /product= "Membrane attractin-2"
XX
PN WO200015651-A1.
XX
PD 23-MAR-2000.
XX
PF 14-SEP-1999; 99WO-US020948.
XX
PR 14-SEP-1998; 98US-0100137P.
XX
PA (DAND) DANA FARBER CANCER INST INC.
XX
PI Duke-Cohan JS, Schlossman SF;

XX
DR WPI; 2000-271373/23.
DR P-PSDB; AAY70691.
XX
PT Isolated nucleic acids encoding human attractin polypeptides useful for
PT enhancing immune responses.
XX
PS Claim 3; Fig 14; 120pp; English.
XX
CC The patent discloses four forms of human attractin polypeptides which
CC enhance immune response by promoting macrophage and monocyte spreading in
CC the presence of T cells. These include soluble attractin-1 and -2 and
CC membrane attractin-1 and -2. These various forms of attractin are encoded
CC by alternatively spliced mRNA molecule transcribed from a single gene.
CC The present sequence is a cDNA encoding human membrane attractin-2.
CC Membrane attractin differs from soluble attractin in having a
CC transmembrane domain and a cytoplasmic domain at the C-terminal.
CC Attractin-2 differs from attractin-1 in having a 74 amino acid insert in
CC the N-terminal. This sequence can be used to enhance immune response in
CC immunosuppressed patients such as those undergoing chemo- and radio-
CC therapy treatment for cancer or those suffering from common variable
CC immunodeficiency syndrome. The proteins may also be used to screen
CC modulators (agonists and antagonists) of immune response which may also
CC be used to regulate immune reactions. Attractin antibodies can be used to
CC inhibit immune response in transplant recipients or patients afflicted
CC with autoimmune disease
XX
SQ Sequence 4290 BP; 1060 A; 1025 C; 1120 G; 1085 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	4290
Score:	7915.00	Matches:	1429
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	3	Gaps:	0

US-09-787-097-12 (1-1429) x AAZ52276 (1-4290)

```
Qy          1 MetValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAla  20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA  60

Qy         21 AlaLeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArg  40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGGGACTGGGACGTGACCAGGGCTGGGAGG 120

Qy         41 ProGlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSerProProLeuArgProArg  60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGGGGCTGGGGGCCGGGCTGCGCCTCCCGCGGCTGCTGTCTCCACCGCTGCGGCCACGG 180

Qy         61 LeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeuProCysGlu  80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGCTGCTGCTGCTGTTGTTGCTCCCGCCGCCGCTGTTGCTGCTGCTGCTGCCCTGTGAG 240

Qy         81 AlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCGAGGCCGCGGCGGCGGCGGCGGCGGTGTCGGGCTCAGCCGCAGCCGAGGCCAAGGAA 300

Qy        101 CysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCys 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGTGACCGGCCCTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGC 360

Qy        121 ProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGly 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGA 420

Qy        141 SerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrp 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGG 480

Qy        161 LeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        481 CTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAG 540

Qy        181 CysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAla 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCT 600

Qy        201 AlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGluValVal 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTT 660

Qy        221 AlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGly 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGA 720

Qy        241 PheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCys 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGT 780

Qy        261 LysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsnTrpLysGlyGlu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAGATCAGTAATAGCAGCGAAACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAA 840

Qy        281 AlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsn 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAAT 900

Qy        301 SerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerVal 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTT 960

Qy        321 ProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuPro 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCC 1020

Qy        341 ArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMet 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATG 1080

Qy        361 PheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeu 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTT 1140

Qy        381 ProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATAC 1200

Qy        401 LysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsnValThrAsnGlu 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AAGGATAAAATTTACATGTATGGAGGAAAAATTGATCCAACTGGGAATGTGACCAATGAG 1260

Qy        421 LeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGlu 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGGAG 1320

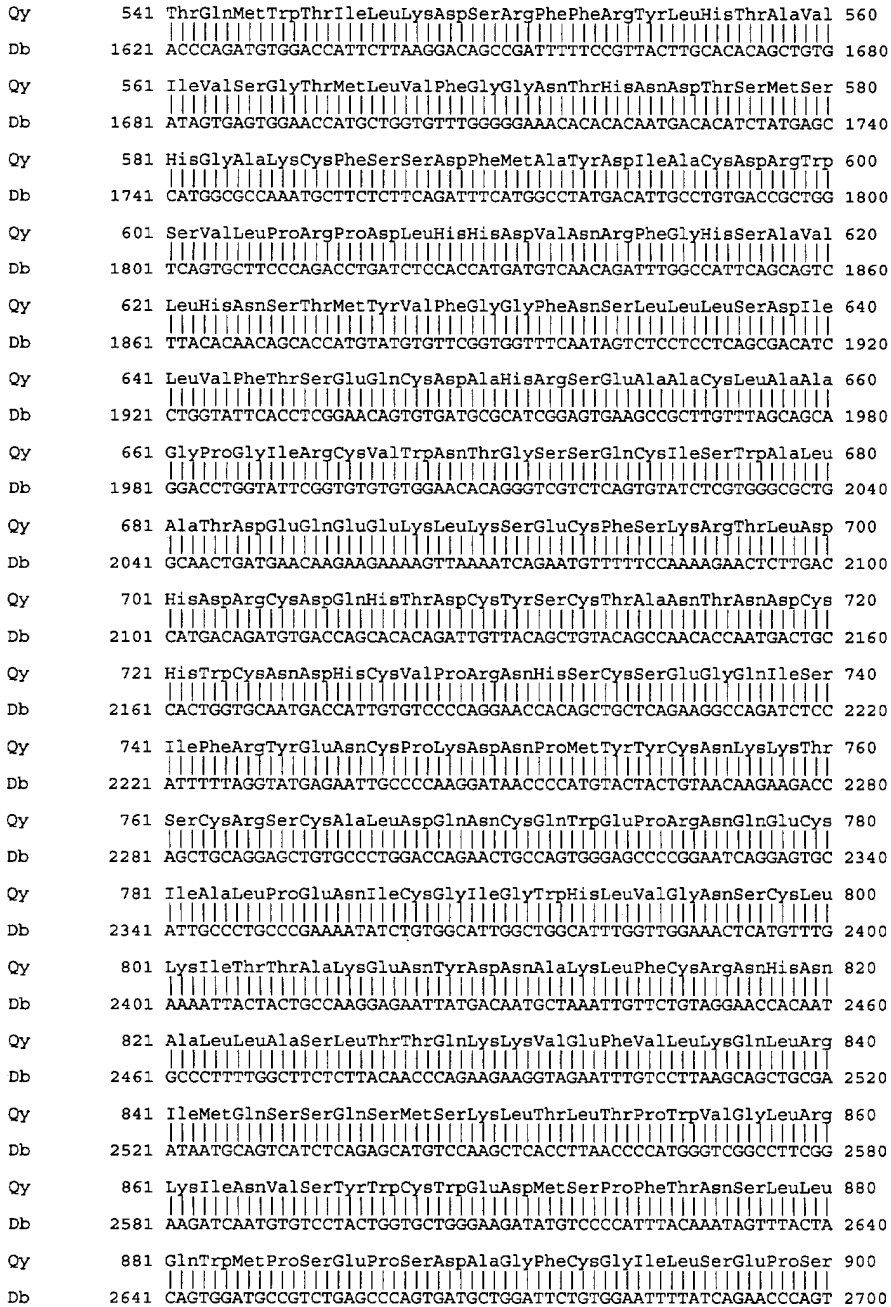
Qy        441 GlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValVal 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTC 1380

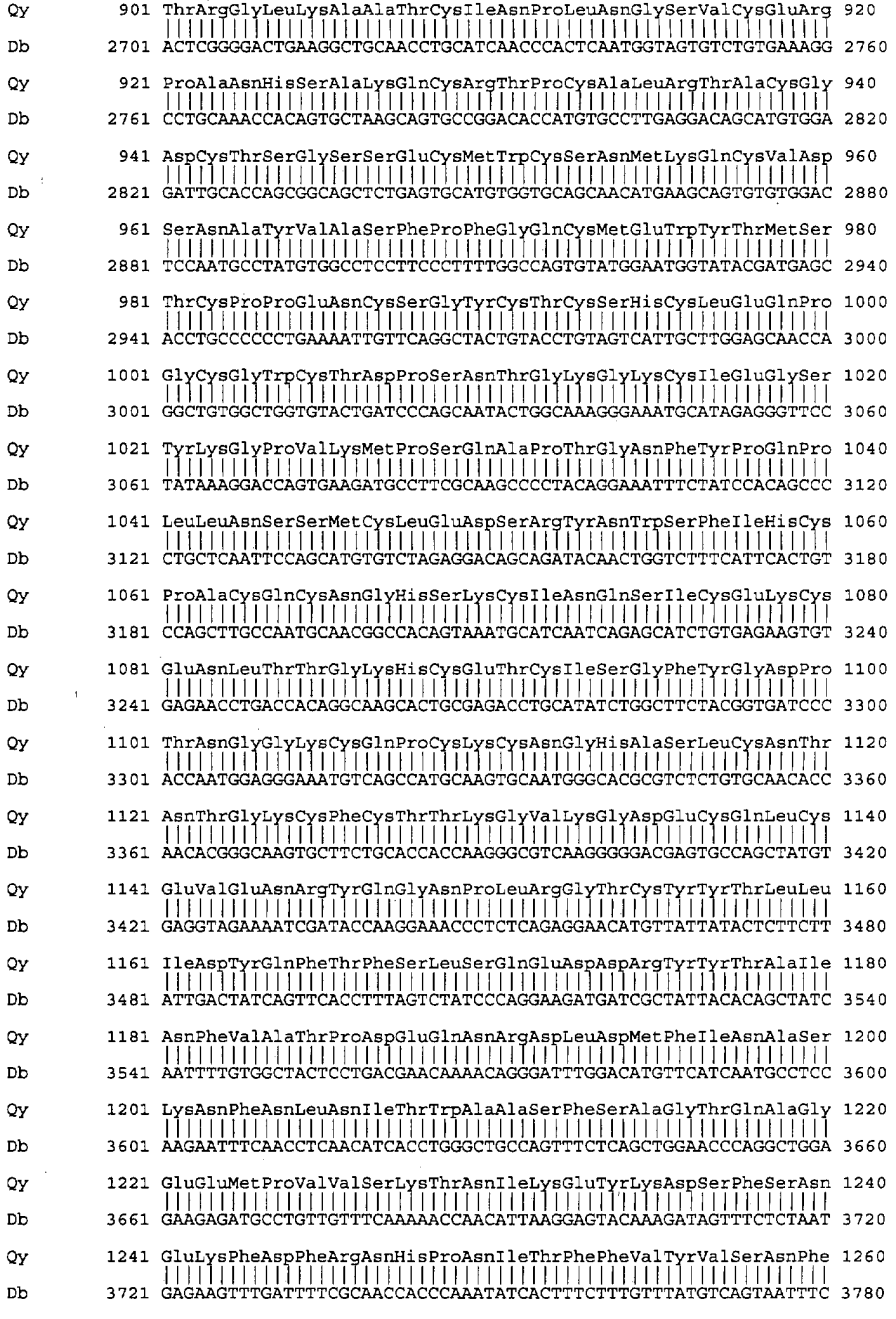
Qy        461 MetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyr 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATAT 1440

Qy        481 AspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGly 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGT 1500

Qy        501 TyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyr 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTAC 1560

Qy        521 LysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAsp 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGAT 1620
```


Qy 541 ThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaVal 560
Db 1621 ACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTG 1680
Qy 561 IleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSer 580
Db 1681 ATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGC 1740
Qy 581 HisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrp 600
Db 1741 CATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGG 1800
Qy 601 SerValLeuProArgProAspLeuHisHisAspValAsnArgPheGlyHisSerAlaVal 620
Db 1801 TCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTC 1860
Qy 621 LeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIle 640
Db 1861 TTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATC 1920
Qy 641 LeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAla 660
Db 1921 CTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCA 1980
Qy 661 GlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeu 680
Db 1981 GGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTG 2040
Qy 681 AlaThrAspGluGlnGluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAsp 700
Db 2041 GCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGAC 2100
Qy 701 HisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCys 720
Db 2101 CATGACAGATGTGACCAGCACACAGATTGTTACAGCTGTACAGCCAACACCAATGACTGC 2160
Qy 721 HisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSer 740
Db 2161 CACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCC 2220
Qy 741 IlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThr 760
Db 2221 ATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCCATGTACTACTGTAACAAGAAGACC 2280
Qy 761 SerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCys 780
Db 2281 AGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGC 2340
Qy 781 IleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeu 800
Db 2341 ATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTG 2400
Qy 801 LysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsn 820
Db 2401 AAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAAT 2460
Qy 821 AlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArg 840
Db 2461 GCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGA 2520
Qy 841 IleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArg 860
Db 2521 ATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGG 2580
Qy 861 LysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeu 880
Db 2581 AAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTA 2640
Qy 881 GlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSer 900
Db 2641 CAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGT 2700


Qy 901 ThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArg 920
Db 2701 ACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGG 2760
Qy 921 ProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGly 940
Db 2761 CCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGA 2820
Qy 941 AspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAsp 960
Db 2821 GATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGAC 2880
Qy 961 SerAsnAlaTyrValAlaSerProProGlyGlnCysMetGluTrpTyrThrMetSer 980
Db 2881 TCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGC 2940
Qy 981 ThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPro 1000
Db 2941 ACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCA 3000
Qy 1001 GlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySer 1020
Db 3001 GGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCC 3060
Qy 1021 TyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnPro 1040
Db 3061 TATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCC 3120
Qy 1041 LeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCys 1060
Db 3121 CTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGT 3180
Qy 1061 ProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCys 1080
Db 3181 CCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGT 3240
Qy 1081 GluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspPro 1100
Db 3241 GAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCC 3300
Qy 1101 ThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThr 1120
Db 3301 ACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACC 3360
Qy 1121 AsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCys 1140
Db 3361 AACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGT 3420
Qy 1141 GluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeu 1160
Db 3421 GAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTT 3480
Qy 1161 IleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIle 1180
Db 3481 ATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATC 3540
Qy 1181 AsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSer 1200
Db 3541 AATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCC 3600
Qy 1201 LysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGly 1220
Db 3601 AAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGA 3660
Qy 1221 GluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsn 1240
Db 3661 GAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAAT 3720
Qy 1241 GluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPhe 1260
Db 3721 GAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTC 3780
Qy 1261 ThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuVal 1280

```
Db       3781 ACCTGGCCCATCAAAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCTGGTA 3840

Qy       1281 GlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuLeuValAlaAlaValVal 1300

Db       3841 CAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGTGGTT 3900

Qy       1301 TrpLysIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeuLeuArgGluMet 1320

Db       3901 TGGAAGATCAAACAAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGAGATG 3960

Qy       1321 GlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThrAspGluGlu 1340

Db       3961 CAACAGATGGCCAGCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGAGGAG 4020

Qy       1341 ProProAspLeuIleGlyGlySerIleLysThrValProLysProIleAlaLeuGluPro 1360

Db       4021 CCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGAGCCG 4080

Qy       1361 CysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGly 1380

Db       4081 TGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCTGGGT 4140

Qy       1381 GlyIleProProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSer 1400

Db       4141 GGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACATTTCT 4200

Qy       1401 GlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAsnArgLysGlnGln 1420

Db       4201 CAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCAGCAG 4260

Qy       1421 ProProAlaGlnProGlyThrCysIle 1429

Db       4261 CCCCCTGCACAGCCTGGGACCTGCATC 4287

RESULT 2
AAS72659
ID   AAS72659 standard; cDNA; 8561 BP.
XX
AC   AAS72659;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #8463.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG08472.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 8463; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 8561 BP; 2246 A; 1983 C; 2065 G; 2267 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                0          Length:          8561
Score:                    7915.00    Matches:         1429
Percent Similarity:       100.00%    Conservative:    0
Best Local Similarity:    100.00%    Mismatches:      0
Query Match:              100.00%    Indels:          0
DB:                       5          Gaps:            0

US-09-787-097-12 (1-1429) x AAS72659 (1-8561)

Qy          1 MetValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAla 20

Db          1 ATGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 60

Qy         21 AlaLeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArg 40

Db         61 GCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGGGACTGGGACGTGACCAGGGCTGGGAGG 120

Qy         41 ProGlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSerProProLeuArgProArg 60

Db        121 CCGGGGCTGGGGGCCGGGCTGCGCCTCCCGCGGCTGCTGTCTCCACCGCTGCGGCCACGG 180

Qy         61 LeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeuProCysGlu 80

Db        181 CTGCTGCTGCTGCTGTTGTTGCTCCCGCCGCCGCTGTTGCTGCTGCTGCTGCCCTGTGAG 240

Qy         81 AlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGlu 100

Db        241 GCCGAGGCCGCGGCGGCGGCGGCGGCGGTGTCGGGCTCAGCCGCAGCCGAGGCCAAGGAA 300

Qy        101 CysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCys 120

Db        301 TGTGACCGGCCCTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGC 360

Qy        121 ProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGly 140

Db        361 CCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGA 420

Qy        141 SerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrp 160

Db        421 TCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGG 480

Qy        161 LeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGlu 180

Db        481 CTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAG 540

Qy        181 CysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAla 200

Db        541 TGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCT 600
```

```
Qy       201 AlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGluValVal 220
Db       601 GCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTT 660

Qy       221 AlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGly 240
Db       661 GCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGA 720

Qy       241 PheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCys 260
Db       721 TTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGT 780

Qy       261 LysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsnTrpLysGlyGlu 280
Db       781 AAGATCAGTAATAGCAGCGAAACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAA 840

Qy       281 AlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsn 300
Db       841 GCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAAT 900

Qy       301 SerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerVal 320
Db       901 TCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTT 960

Qy       321 ProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuPro 340
Db       961 CCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCC 1020

Qy       341 ArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMet 360
Db      1021 AGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATG 1080

Qy       361 PheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeu 380
Db      1081 TTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTT 1140

Qy       381 ProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyr 400
Db      1141 CCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATAC 1200

Qy       401 LysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsnValThrAsnGlu 420
Db      1201 AAGGATAAAATTTACATGTATGGAGGAAAAATTGATCCAACTGGGAATGTGACCAATGAG 1260

Qy       421 LeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGlu 440
Db      1261 TTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGGAG 1320

Qy       441 GlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValVal 460
Db      1321 CAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTC 1380

Qy       461 MetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyr 480
Db      1381 ATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATAT 1440

Qy       481 AspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGly 500
Db      1441 GATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGT 1500

Qy       501 TyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyr 520
Db      1501 TACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTAC 1560

Qy       521 LysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAsp 540
Db      1561 AAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGAT 1620

Qy       541 ThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaVal 560
Db      1621 ACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTG 1680

Qy       561 IleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSer 580
Db      1681 ATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGC 1740

Qy       581 HisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrp 600
Db      1741 CATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGG 1800

Qy       601 SerValLeuProArgProAspLeuHisHisAspValAsnArgPheGlyHisSerAlaVal 620
Db      1801 TCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTC 1860

Qy       621 LeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIle 640
Db      1861 TTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATC 1920

Qy       641 LeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAla 660
Db      1921 CTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCA 1980

Qy       661 GlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeu 680
Db      1981 GGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTG 2040

Qy       681 AlaThrAspGluGlnGluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAsp 700
Db      2041 GCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGAC 2100

Qy       701 HisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCys 720
Db      2101 CATGACAGATGTGACCAGCACACAGATTGTTACAGCTGTACAGCCAACACCAATGACTGC 2160

Qy       721 HisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSer 740
Db      2161 CACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCC 2220

Qy       741 IlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThr 760
Db      2221 ATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCCATGTACTACTGTAACAAGAAGACC 2280

Qy       761 SerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCys 780
Db      2281 AGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGC 2340

Qy       781 IleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeu 800
Db      2341 ATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTG 2400

Qy       801 LysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsn 820
Db      2401 AAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAAT 2460

Qy       821 AlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArg 840
Db      2461 GCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGA 2520

Qy       841 IleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArg 860
Db      2521 ATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGG 2580

Qy       861 LysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeu 880
Db      2581 AAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTA 2640

Qy       881 GlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSer 900
Db      2641 CAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGT 2700

Qy       901 ThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArg 920
Db      2701 ACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGG 2760

Qy       921 ProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGly 940
```

```
Db       2761 CCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGA 2820
Qy        941 AspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAsp 960
Db       2821 GATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGAC 2880
Qy        961 SerAsnAlaTyrValAlaSerProProGlyGlnCysMetGluTrpTyrThrMetSer 980
Db       2881 TCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGC 2940
Qy        981 ThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPro 1000
Db       2941 ACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCA 3000
Qy       1001 GlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySer 1020
Db       3001 GGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCC 3060
Qy       1021 TyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnPro 1040
Db       3061 TATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCC 3120
Qy       1041 LeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCys 1060
Db       3121 CTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGT 3180
Qy       1061 ProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCys 1080
Db       3181 CCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGT 3240
Qy       1081 GluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspPro 1100
Db       3241 GAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCC 3300
Qy       1101 ThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThr 1120
Db       3301 ACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACC 3360
Qy       1121 AsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCys 1140
Db       3361 AACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGT 3420
Qy       1141 GluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeu 1160
Db       3421 GAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTT 3480
Qy       1161 IleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIle 1180
Db       3481 ATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATC 3540
Qy       1181 AsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSer 1200
Db       3541 AATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCC 3600
Qy       1201 LysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGly 1220
Db       3601 AAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGA 3660
Qy       1221 GluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsn 1240
Db       3661 GAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAAT 3720
Qy       1241 GluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPhe 1260
Db       3721 GAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTC 3780
Qy       1261 ThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuVal 1280
Db       3781 ACCTGGCCCATCAAAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCTGGTA 3840
Qy       1281 GlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuLeuValAlaAlaValVal 1300
```

```
Db       3841 CAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGTGGTT 3900
Qy       1301 TrpLysIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeuLeuArgGluMet 1320
Db       3901 TGGAAGATCAAACAAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGAGATG 3960
Qy       1321 GlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThrAspGluGlu 1340
Db       3961 CAACAGATGGCCAGCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGAGGAG 4020
Qy       1341 ProProAspLeuIleGlyGlySerIleLysThrValProLysProIleAlaLeuGluPro 1360
Db       4021 CCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGAGCCG 4080
Qy       1361 CysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGly 1380
Db       4081 TGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCTGGGT 4140
Qy       1381 GlyIleProProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSer 1400
Db       4141 GGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACATTTCT 4200
Qy       1401 GlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAsnArgLysGlnGln 1420
Db       4201 CAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCAGCAG 4260
Qy       1421 ProProAlaGlnProGlyThrCysIle 1429
Db       4261 CCCCCTGCACAGCCTGGGACCTGCATC 4287
```

RESULT 3
AAZ91920
ID AAZ91920 standard; cDNA; 8590 BP.
XX
AC AAZ91920;
XX
DT 08-JUN-2000 (first entry)
XX
DE Human mahogany protein coding sequence #2.
XX
KW Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;
KW weight regulation; cell therapy; body weight disorder; cachexia;
KW anorexia; hyperpigmentation; increased metabolic rate disorder;
KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.
XX
OS Homo sapiens.
XX
PN WO200005373-A2.
XX
PD 03-FEB-2000.
XX
PF 21-JUL-1999; 99WO-US016484.
XX
PR 21-JUL-1998; 98US-0093630P.
PR 20-OCT-1998; 98US-0104978P.
PR 05-FEB-1999; 99US-00245041.
XX
PA (MILL-) MILLENIUM PHARM INC.
XX
PI Moore K, Nagle DL;
XX
DR WPI; 2000-195103/17.
DR P-PSDB; AAY81807.
XX
PT New human and murine mahogany genes, useful, e.g. for diagnosis and
PT treatment of body weight disorders.
XX
PS Claim 1; Fig 18a; 188pp; English.
XX
CC This sequence represents a human mahogany gene of the invention. The
CC mahogany genes are used: (i) to produce recombinant mahogany (mg)
CC proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming
CC therapeutics; (iii) as a source of diagnostic probes and primers for

CC detecting expression of mg genes or mutations, regulatory defects, in
CC this gene, or for isolation of related sequences; and (iv) in (cell-
CC based) gene therapy. (II) are used to raise specific antibodies (Ab); to
CC identify other (extra)cellular products involved in weight regulation,
CC and to screen for agents that disrupt interaction between (II) and other
CC macromolecules. The Ab are used to detect abnormal levels (or function)
CC of (II) (for diagnosis, prognosis or monitoring of treatment); to
CC evaluate (II)-expressing cells intended for cell therapy, and as
CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the
CC mg polypeptide) are used to identify agents (A) that modulate mg
CC activity. (A) are potentially useful for the treatment of body weight
CC disorders, particularly obesity, cachexia or anorexia, or other
CC conditions associated with the mg gene such as hyperpigmentation,
CC hyperphagia and disorders that result in increased metabolic rate
XX
SQ Sequence 8590 BP; 2275 A; 1990 C; 2061 G; 2264 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	8590
Score:	7696.50	Matches:	1401
Percent Similarity:	97.91%	Conservative:	2
Best Local Similarity:	97.77%	Mismatches:	17
Query Match:	97.24%	Indels:	13
DB:	3	Gaps:	3

US-09-787-097-12 (1-1429) x AAZ91920 (1-8590)

```
Qy          2 ValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAlaAla 21
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         10 GTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCG 69

Qy         22 LeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArgPro 41
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 CTCGCGGGCAGGAGCGGCGGGCCGCACCGACCCTGCACCGCGACAGGGGCCTGGAGGCCG 129

Qy         42 GlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSer---------------ProPro 56
              |||   |||  |||   |||||||||:::||||||                |||||
Db        130 GGACCGCGCGCCCGGCTGTGTCTCCCGCGGGTGCTGTCGCGGGCGCTGCCCCCGCCGCCG 189

Qy         57 LeuArgProArgLeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeu 76
              |||  |||  ||||||     ||||||||||||||||
Db        190 CTGCTGCCG---CTGCTCTTTTCGCTGCTGCTGCCG---------------------- 225

Qy         77 LeuProCysGluAlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAla 96
              ||||||  ||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db        226 CTGCCCCGGGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCC 285

Qy         97 GluAlaLysGluCysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGly 116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 GAGGCCAAGGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGC 345

Qy        117 GlnCysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPhe 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 CAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTC 405

Qy        137 ArgLeuThrGlySerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThr 156
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        406 AGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACG 465

Qy        157 LysCysThrTrpLeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHis 176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        466 AAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCAT 525

Qy        177 PheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAla 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        526 TTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCA 585

Qy        197 ProLeuValAlaAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrVal 216
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        586 CCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTC 645

Qy        217 ProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyr 236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        646 CCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTAT 705

Qy        237 AsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGly 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        706 AATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGC 765

Qy        257 ArgGlyGluCysLysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsn 276
              ||||||||||||||||||||||||||||||:::||||||||||||||||||||||||||||
Db        766 CGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAAC 825

Qy        277 TrpLysGlyGluAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArg 296
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        826 TGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGA 885

Qy        297 GlyIleCysAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyPro 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        886 GGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCT 945

Qy        317 GlyCysSerValProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsn 336
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        946 GGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAAC 1005

Qy        337 LeuLysLeuProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValVal 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1006 TTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTT 1065

Qy        357 GlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSer 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1066 GGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCT 1125

Qy        377 ArgGluTrpLeuProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSer 396
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1126 AGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCT 1185

Qy        397 LeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsn 416
              ||||||||||||||||||||||||||||||||||||||||||||||||||||  ||||||||
Db       1186 TTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAAT 1245

Qy        417 ValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrPro 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1246 GTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCT 1305

Qy        437 LysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsn 456
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1306 AAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAAT 1365

Qy        457 GlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsn 476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1366 GGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAAT 1425

Qy        477 ValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeu 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1426 GTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTT 1485

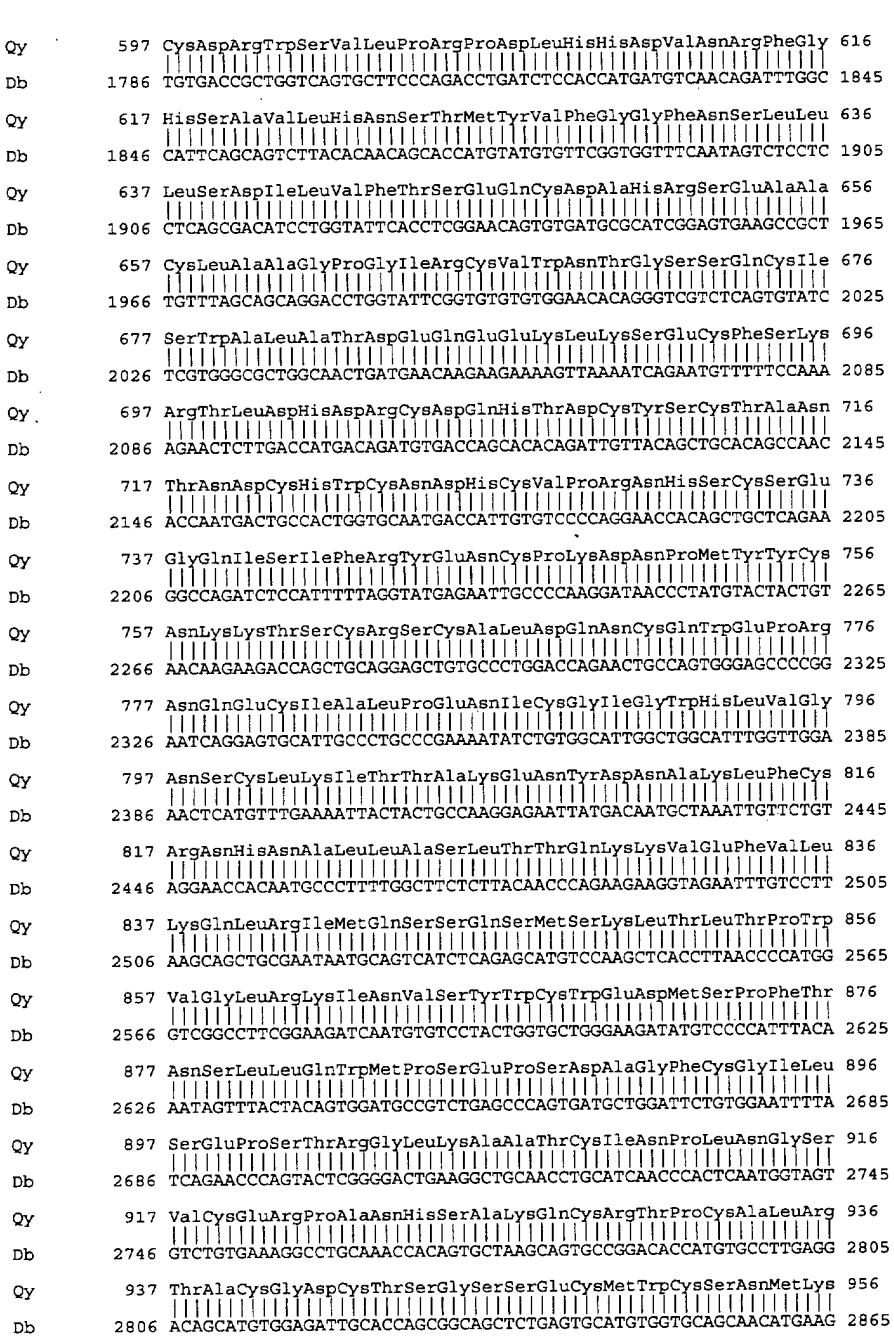
Qy        497 ValGlnGlyGlyTyrGlyHisSerServalTyrAspHisArgThrArgAlaLeuTyrVal 516
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1486 GTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTT 1545

Qy        517 HisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArg 536
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1546 CATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGA 1605

Qy        537 TyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeu 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1606 TATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTG 1665

Qy        557 HisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAsp 576
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1666 CACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGAC 1725

Qy        577 ThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAla 596
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1726 ACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCC 1785
```


Qy 597 CysAspArgTrpSerValLeuProArgProAspLeuHisHisAspValAsnArgPheGly 616
Db 1786 TGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGC 1845
Qy 617 HisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeu 636
Db 1846 CATTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTC 1905
Qy 637 LeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAla 656
Db 1906 CTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCT 1965
Qy 657 CysLeuAlaAlaGlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIle 676
Db 1966 TGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATC 2025
Qy 677 SerTrpAlaLeuAlaThrAspGluGlnGluGluLysLeuLysSerGluCysPheSerLys 696
Db 2026 TCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAA 2085
Qy 697 ArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsn 716
Db 2086 AGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAAC 2145
Qy 717 ThrAsnAspCysHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGlu 736
Db 2146 ACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAA 2205
Qy 737 GlyGlnIleSerIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCys 756
Db 2206 GGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGT 2265
Qy 757 AsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArg 776
Db 2266 AACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGG 2325
Qy 777 AsnGlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGly 796
Db 2326 AATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGA 2385
Qy 797 AsnSerCysLeuLysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCys 816
Db 2386 AACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGT 2445
Qy 817 ArgAsnHisAsnAlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeu 836
Db 2446 AGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTT 2505
Qy 837 LysGlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrp 856
Db 2506 AAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGG 2565
Qy 857 ValGlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThr 876
Db 2566 GTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACA 2625
Qy 877 AsnSerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeu 896
Db 2626 AATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTA 2685
Qy 897 SerGluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySer 916
Db 2686 TCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGT 2745
Qy 917 ValCysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArg 936
Db 2746 GTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGG 2805
Qy 937 ThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLys 956
Db 2806 ACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAG 2865
Qy 957 GlnCysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrp 976
Db 2866 CAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGG 2925
Qy 977 TyrThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCys 996
Db 2926 TATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGC 2985
Qy 997 LeuGluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCys 1016
Db 2986 TTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGC 3045
Qy 1017 IleGluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPhe 1036
Db 3046 ATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTC 3105
Qy 1037 TyrProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSer 1056
Db 3106 TATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCT 3165
Qy 1057 PheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIle 1076
Db 3166 TTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATC 3225
Qy 1077 CysGluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPhe 1096
Db 3226 TGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTC 3285
Qy 1097 TyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSer 1116
Db 3286 TACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCT 3345
Qy 1117 LeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGlu 1136
Db 3346 CTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAG 3405
Qy 1137 CysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyr 1156
Db 3406 TGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTAT 3465
Qy 1157 TyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyr 1176
Db 3466 TATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTAT 3525
Qy 1177 TyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPhe 1196
Db 3526 TACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTC 3585
Qy 1197 IleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGly 1216
Db 3586 ATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGA 3645
Qy 1217 ThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAsp 1236
Db 3646 ACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGAT 3705
Qy 1237 SerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyr 1256
Db 3706 AGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTAT 3765
Qy 1257 ValSerAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPhe 1276
Db 3766 GTCAGTAATTTCACCTGGCCCATCAAAATTCAGATTGCCTTCTCTCAGCACAGCAATTTT 3825
Qy 1277 MetAspLeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuLeuVal 1296
Db 3826 ATGGACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTG 3885
Qy 1297 AlaAlaValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeu 1316
Db 3886 GCTGCTGTGGTTTGGAAGATCAAACAAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTT 3945
Qy 1317 LeuArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGlu 1336

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3946 CTTCGAGAGATGCAACAGATGGCCAGCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAA 4005

Qy       1337 ThrAspGluGluProProAspLeuIleGlyGlySerIleLysThrValProLysProIle 1356
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4006 ACAGATGAGGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATT 4065

Qy       1357 AlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuPro 1376
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4066 GCACTGGAGCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCT 4125

Qy       1377 ArgGlyLeuGlyGlyIleProProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeu 1396
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4126 CGAGGCCTGGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTG 4185

Qy       1397 ValAspIleSerGlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAsn 1416
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4186 GTGGACATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAAC 4245

Qy       1417 ArgLysGlnGlnProProAlaGlnProGlyThrCysIle 1429
                ||||||||||||||||||||||||||||||||||||||
Db       4246 CGGAAGCAGCAGCCCCCTGCACAGCCTGGGACCTGCATC 4284
```

RESULT 4
AAZ52275
ID AAZ52275 standard; cDNA; 4068 BP.
XX
AC AAZ52275;
XX
DT 18-JUL-2000 (first entry)
XX
DE Human membrane attractin-1 cDNA.
XX
KW Human; membrane attractin-1; immune response; macrophage; monocyte;
KW T cell; immunostimulant; immunosuppressed patient; cancer;
KW immunodeficiency syndrome; transplant; autoimmune disease; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1. .4068
FT /*tag= a
FT /product= "Membrane attractin-1"
XX
PN WO200015651-A1.
XX
PD 23-MAR-2000.
XX
PF 14-SEP-1999; 99WO-US020948.
XX
PR 14-SEP-1998; 98US-0100137P.
XX
PA (DAND) DANA FARBER CANCER INST INC.
XX
PI Duke-Cohan JS, Schlossman SF;
XX
DR WPI; 2000-271373/23.
DR P-PSDB; AAY70690.
XX
PT Isolated nucleic acids encoding human attractin polypeptides useful for
PT enhancing immune responses.
XX
PS Claim 3; Fig 10; 120pp; English.
XX
CC The patent discloses four forms of human attractin polypeptides which
CC enhance immune response by promoting macrophage and monocyte spreading in
CC the presence of T cells. These include soluble attractin-1 and -2 and
CC membrane attractin-1 and -2. These various forms of attractin are encoded
CC by alternatively spliced mRNA molecule transcribed from a single gene.
CC The present sequence is a cDNA encoding human membrane attractin-1.
CC Membrane attractin differs from soluble attractin in having a
CC transmembrane domain and a cytoplasmic domain at the C-terminal. This
CC sequence can be used to enhance immune response in immunosuppressed
CC patients such as those undergoing chemo- and radio-therapy treatment for
CC cancer or those suffering from common variable immunodeficiency syndrome.
CC The proteins may also be used to screen modulators (agonists and
CC antagonists) of immune response which may also be used to regulate immune
CC reactions. Attractin antibodies can be used to inhibit immune response in
CC transplant recipients or patients afflicted with autoimmune disease
XX
SQ Sequence 4068 BP; 1042 A; 948 C; 1027 G; 1051 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	4068
Score:	7478.00	Matches:	1353
Percent Similarity:	94.75%	Conservative:	1
Best Local Similarity:	94.68%	Mismatches:	1
Query Match:	94.48%	Indels:	74
DB:	3	Gaps:	1

US-09-787-097-12 (1-1429) x AAZ52275 (1-4068)

```
Qy          1 MetValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 60

Qy         21 AlaLeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArg 40
              ||||||||||||||||||||||||||||||||
Db         61 GCGCTCGCGGGCAGGAGCGGCGGGCCGCAC---------------------------------- 90

Qy         41 ProGlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSerProProLeuArgProArg 60

Db         90 ---------------------------------------------------------------- 90

Qy         61 LeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeuProCysGlu 80

Db         90 ---------------------------------------------------------------- 90

Qy         81 AlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGlu 100

Db         90 ---------------------------------------------------------------- 90

Qy        101 CysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCys 120
                          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 -----------TGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGC 138

Qy        121 ProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGly 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 CCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGA 198

Qy        141 SerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrp 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 TCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGG 258

Qy        161 LeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 CTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAG 318

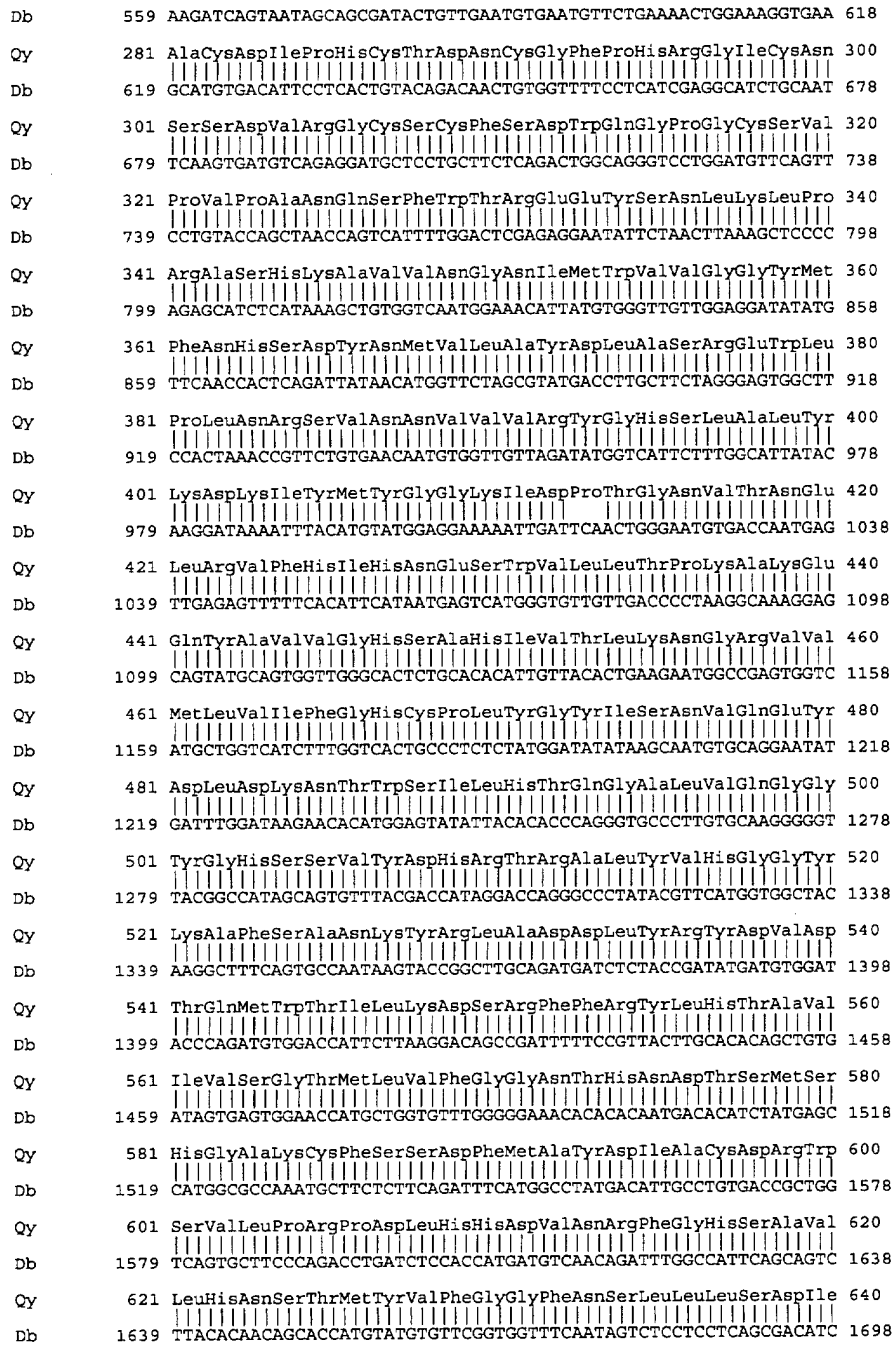
Qy        181 CysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAla 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 TGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCT 378

Qy        201 AlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGluValVal 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 GCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTT 438

Qy        221 AlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGly 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        439 GCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGA 498

Qy        241 PheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCys 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 TTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGT 558

Qy        261 LysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsnTrpLysGlyGlu 280
              ||||||||||||||||||:::||||||||||||||||||||||||||||||||||||||||
```


Db 559 AAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAA 618
Qy 281 AlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsn 300
Db 619 GCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAAT 678
Qy 301 SerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerVal 320
Db 679 TCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTT 738
Qy 321 ProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuPro 340
Db 739 CCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCC 798
Qy 341 ArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMet 360
Db 799 AGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATG 858
Qy 361 PheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeu 380
Db 859 TTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTT 918
Qy 381 ProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyr 400
Db 919 CCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATAC 978
Qy 401 LysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsnValThrAsnGlu 420
Db 979 AAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAATGAG 1038
Qy 421 LeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGlu 440
Db 1039 TTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGGAG 1098
Qy 441 GlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValVal 460
Db 1099 CAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTC 1158
Qy 461 MetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyr 480
Db 1159 ATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATAT 1218
Qy 481 AspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGly 500
Db 1219 GATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGT 1278
Qy 501 TyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyr 520
Db 1279 TACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTAC 1338
Qy 521 LysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAsp 540
Db 1339 AAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGAT 1398
Qy 541 ThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaVal 560
Db 1399 ACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTG 1458
Qy 561 IleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSer 580
Db 1459 ATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGC 1518
Qy 581 HisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrp 600
Db 1519 CATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGG 1578
Qy 601 SerValLeuProArgProAspLeuHisHisAspValAsnArgPheGlyHisSerAlaVal 620
Db 1579 TCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTC 1638
Qy 621 LeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIle 640
Db 1639 TTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATC 1698
Qy 641 LeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAla 660
Db 1699 CTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCA 1758
Qy 661 GlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeu 680
Db 1759 GGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTG 1818
Qy 681 AlaThrAspGluGlnGluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAsp 700
Db 1819 GCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGAC 1878
Qy 701 HisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCys 720
Db 1879 CATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGACTGC 1938
Qy 721 HisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSer 740
Db 1939 CACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCC 1998
Qy 741 IlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThr 760
Db 1999 ATTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAAGACC 2058
Qy 761 SerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCys 780
Db 2059 AGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGC 2118
Qy 781 IleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeu 800
Db 2119 ATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTG 2178
Qy 801 LysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsn 820
Db 2179 AAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAAT 2238
Qy 821 AlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArg 840
Db 2239 GCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGA 2298
Qy 841 IleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArg 860
Db 2299 ATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGG 2358
Qy 861 LysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeu 880
Db 2359 AAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTA 2418
Qy 881 GlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSer 900
Db 2419 CAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGT 2478
Qy 901 ThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArg 920
Db 2479 ACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGG 2538
Qy 921 ProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGly 940
Db 2539 CCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGA 2598
Qy 941 AspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAsp 960
Db 2599 GATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGAC 2658
Qy 961 SerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSer 980
Db 2659 TCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGC 2718
Qy 981 ThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPro 1000
Db 2719 ACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCA 2778

```
Qy       1001 GlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySer 1020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2779 GGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCC 2838

Qy       1021 TyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnPro 1040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2839 TATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCC 2898

Qy       1041 LeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCys 1060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2899 CTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGT 2958

Qy       1061 ProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCys 1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2959 CCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGT 3018

Qy       1081 GluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspPro 1100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3019 GAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCC 3078

Qy       1101 ThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThr 1120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3079 ACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACC 3138

Qy       1121 AsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCys 1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3139 AACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGT 3198

Qy       1141 GluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeu 1160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3199 GAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTT 3258

Qy       1161 IleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIle 1180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3259 ATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATC 3318

Qy       1181 AsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSer 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3319 AATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCC 3378

Qy       1201 LysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGly 1220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3379 AAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGA 3438

Qy       1221 GluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsn 1240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3439 GAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAAT 3498

Qy       1241 GluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPhe 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3499 GAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTC 3558

Qy       1261 ThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuVal 1280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3559 ACCTGGCCCATCAAAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCTGGTA 3618

Qy       1281 GlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuLeuValAlaAlaValVal 1300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3619 CAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGTGGTT 3678

Qy       1301 TrpLysIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeuLeuArgGluMet 1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3679 TGGAAGATCAAACAAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGAGATG 3738

Qy       1321 GlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThrAspGluGlu 1340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3739 CAACAGATGGCCAGCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGAGGAG 3798

Qy       1341 ProProAspLeuIleGlyGlySerIleLysThrValProLysProIleAlaLeuGluPro 1360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3799 CCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGAGCCG 3858

Qy       1361 CysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGly 1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3859 TGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCTGGGT 3918

Qy       1381 GlyIleProProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSer 1400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3919 GGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACATTTCT 3978

Qy       1401 GlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAsnArgLysGlnGln 1420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3979 CAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCAGCAG 4038

Qy       1421 ProProAlaGlnProGlyThrCysIle 1429
             |||||||||||||||||||||||||||
Db       4039 CCCCCTGCACAGCCTGGGACCTGCATC 4065
```

RESULT 5
ADB53249
ID ADB53249 standard; DNA; 8739 BP.
XX
AC ADB53249;
XX
DT 04-DEC-2003 (first entry)
XX
DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3791.
XX
KW toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KW toxicity marker; toxicity progression; drug screening;
KW primary rat hepatocyte toxicity modelling; gene; ds.
XX
OS Rattus norvegicus.
XX
PN WO2003065993-A2.
XX
PD 14-AUG-2003.
XX
PF 04-FEB-2003; 2003WO-US003482.
XX
PR 04-FEB-2002; 2002US-0353171P.
PR 13-MAR-2002; 2002US-0363534P.
PR 08-APR-2002; 2002US-0370248P.
PR 10-APR-2002; 2002US-0371134P.
PR 10-APR-2002; 2002US-0371135P.
PR 10-APR-2002; 2002US-0371150P.
PR 11-APR-2002; 2002US-0371413P.
PR 19-APR-2002; 2002US-0373601P.
PR 19-APR-2002; 2002US-0373602P.
PR 22-APR-2002; 2002US-0374139P.
PR 08-MAY-2002; 2002US-0378370P.
PR 09-MAY-2002; 2002US-0378652P.
PR 09-MAY-2002; 2002US-0378653P.
PR 09-MAY-2002; 2002US-0378665P.
PR 09-JUL-2002; 2002US-0394230P.
PR 09-JUL-2002; 2002US-0394253P.
PR 04-SEP-2002; 2002US-0407688P.
PR 28-JAN-2003; 2003US-0442900P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
PI Elashoff M;
XX
DR WPI; 2003-731472/69.
XX
PT Determining if a compound induces a toxic effect on a tissue or cell, for
PT identifying hepatotoxic compounds, comprises comparing a gene expression
PT profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT mean values.
XX
PS Claim 44; SEQ ID NO 3791; 874pp; English.
XX
CC The present invention describes a method for determining whether a
CC compound induces a toxic effect on a tissue or cell. The method comprises
CC preparing a gene expression profile of a tissue or cell sample exposed to

CC the compound, and comparing the gene expression profile to a database
CC comprising data or information on the Tox mean and non-Tox mean value.
CC The method is useful for predicting or identifying at least one toxic
CC effect, particularly hepatotoxicity, of a test or unknown compound. The
CC genes listed in the specification are useful as diagnostic or toxicity
CC markers for the prediction or identification of the physiological state
CC of tissue or cell sample that has been exposed to a compound, or to
CC identify or predict the toxic effects of a compound or an agent. These
CC may also be used as markers for monitoring toxicity progression or for
CC drug screening. The present sequence represents a primary rat hepatocyte
CC toxicity modelling related gene sequence from the present invention.
XX
SQ Sequence 8739 BP; 2288 A; 2115 C; 2109 G; 2227 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	8739
Score:	7342.50	Matches:	1331
Percent Similarity:	95.25%	Conservative:	33
Best Local Similarity:	92.95%	Mismatches:	65
Query Match:	92.77%	Indels:	3
DB:	9	Gaps:	1

US-09-787-097-12 (1-1429) x ADB53249 (1-8739)

```
Qy          1 MetValAlaAlaAlaAla---------AlaThrGluAlaArgLeuArgArgArgThrAla 17
              ||||||||||||||||||         ||||||||||||||||||||||      |||
Db         61 ATGGTGGCGGCGGCGGCGGCGGCGGAGGCGACTGAGGCGCGGTTGAGGGGGTACACGACG 120

Qy         18 AlaThrAlaAlaLeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArg 37
              ||||||||||||   ||||||                ||| |||          |||
Db        121 GCGACAGCAGCGCCCGCGGGCTGGAAGGAGAGGCAGCACCGACCCTGCGCCGCGACAGGG 180

Qy         38 AlaGlyArgProGlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSerProProLeu 57
              |||  ||||||          |||||||  ||||||||:::||||||       |||
Db        181 GCCTGGAGGCCGTGGCCGCGCGCCGGGCTGTGTCTCCCGCGGGTGCTGTCGCGGGCGCTA 240

Qy         58 ArgProArgLeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeuLeu 77
                 |||      ||||||  ||||||  |||          ||||||||||||  |||
Db        241 TCCCCGCCGCCGCTGCTGCCGCTGCTGCCGCTGCTCTTTTCGCTATTGCTGCTGCCACTG 300

Qy         78 ProCysGluAlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAlaGlu 97
              |||  ||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db        301 CCCCGGGAGGCCGAGGCCGCTGCGGTGGCGGCGGCTGTGTCAGGCTCGGCCGCAGCTGAG 360

Qy         98 AlaLysGluCysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGlyGln 117
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCAAGGAATGTGACCGGCCGTGTGTCAACGGCGGCCGCTGCAACCCTGGCACCGGCCAG 420

Qy        118 CysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPheArg 137
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGCGTCTGCCCCACGGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGA 480

Qy        138 LeuThrGlySerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLys 157
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTAACTGGATCTTCTGGATTTGTAACAGATGGACCTGGGAATTATAAATATAAAACGAAA 540

Qy        158 CysThrTrpLeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPhe 177
              ||||||||||||||||||||||||||:::||||||||||||||||||||||||||||||
Db        541 TGCACATGGCTCATTGAAGGACAGCCAAATAAGATAATGAGACTTCGATTCAACCATTTT 600

Qy        178 AlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaPro 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCTACTGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATTTACGCACCT 660

Qy        198 LeuValAlaAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValPro 217
              |||:::||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        661 CTGATTGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCCCCT 720

Qy        218 GluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsn 237
              ||||||      ||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAAGTCACTGTCACTTCAGGTTATGCACTACTGCATTTTTTCAGTGATGCTGCTTATAAT 780

Qy        238 LeuThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArg 257
              ||||||||||||||||||||||||||||||:::|||||||||||||||||||||||||||
Db        781 CTGACTGGATTTAATATCACTTACAATTTTGACATGTGCCCGAATAATTGCTCAGGCCGA 840

Qy        258 GlyGluCysLysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsnTrp 277
              ||||||||||    ||||||||||    ||||||||||||||||||||||||||||||||
Db        841 GGAGAATGTAAGAGCAGTAACAGCAGCAGCACTGTTGAGTGTGAATGTTCTGAAAACTGG 900

Qy        278 LysGlyGluAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGly 297
              |||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAAGGGGAGTCCTGTGACATTCCTCACTGTACAGACAACTGTGGCTTTCCTCACCGAGGC 960

Qy        298 IleCysAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyProGly 317
              ||||||||||:::||||||   ||||||||||||||||||      ||||||||||||||
Db        961 ATCTGTAACGCAAGTGACACCAGAGGGTGCTCCTGCTTCCCTCACTGGCAGGGTCCTGGA 1020

Qy        318 CysSerValProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeu 337
              ||||||:::||||||||||||||||||||||||||||||||||||||||||||:::||||
Db       1021 TGTTCAATTCCTGTGCCAGCTAACCAGTCTTTTTGGACTCGAGAAGAATATTCTGACTTA 1080

Qy        338 LysLeuProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValValGly 357
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGCTTCCCAGAGCCTCTCATAAAGCTGTGGTCAATGGAAATATAATGTGGGTTGTTGGC 1140

Qy        358 GlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArg 377
              ||||||||||||||||||||||||||:::||||||||||||||||||||||||||||||
Db       1141 GGGTATATGTTCAACCATTCAGATTACAGCATGGTTTTAGCGTATGACCTGGCTTCTAGG 1200

Qy        378 GluTrpLeuProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSerLeu 397
              |||||||||  ||||||.  |||||||||:::||||||||||||||||||||||||||||
Db       1201 GAATGGCTTTCACTAAACCATTCTGTGAACAGTGTGGTTGTAAGATATGGTCATTCTTTG 1260

Qy        398 AlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsnVal 417
              |||||:::||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GCATTACATAAGGATAAAATCTACATGTATGGAGGAAAAATTGATTCAACAGGGAATGTG 1320

Qy        418 ThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrProLys 437
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ACCAACGAGCTGAGAGTATTTCATATTCATAATGAATCATGGGTATTGTTAACCCCCAAA 1380

Qy        438 AlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsnGly 457
              ||||||:::||||||||||||||||||||||||||||||||||||||||||||:::|||
Db       1381 GCTAAGGATCAGTATGCAGTGGTTGGACACTCAGCACACATTGTTACACTGTCATCTGGC 1440

Qy        458 ArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsnVal 477
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CGTGTGGTCATGTTGGTCATCTTCGGTCATTGTCCGCTCTACGGATATATAAGTGTTGTG 1500

Qy        478 GlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuVal 497
              ||||||||||||||||:::||||||||||||||||||||||||||||||||||||||||
Db       1501 CAGGAATATGACTTGGAAAAGAACACATGGAGTATATTACAAACTCAAGGTGCTCTAGTG 1560

Qy        498 GlnGlyGlyTyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrValHis 517
              ||||||||||||||||||||||||||||||||||||||||||||||||:::||||||||||
Db       1561 CAAGGGGGTTATGGCCACAGCAGTGTTTATGATCACAGGACCAAGGCTCTGTATGTTCAT 1620

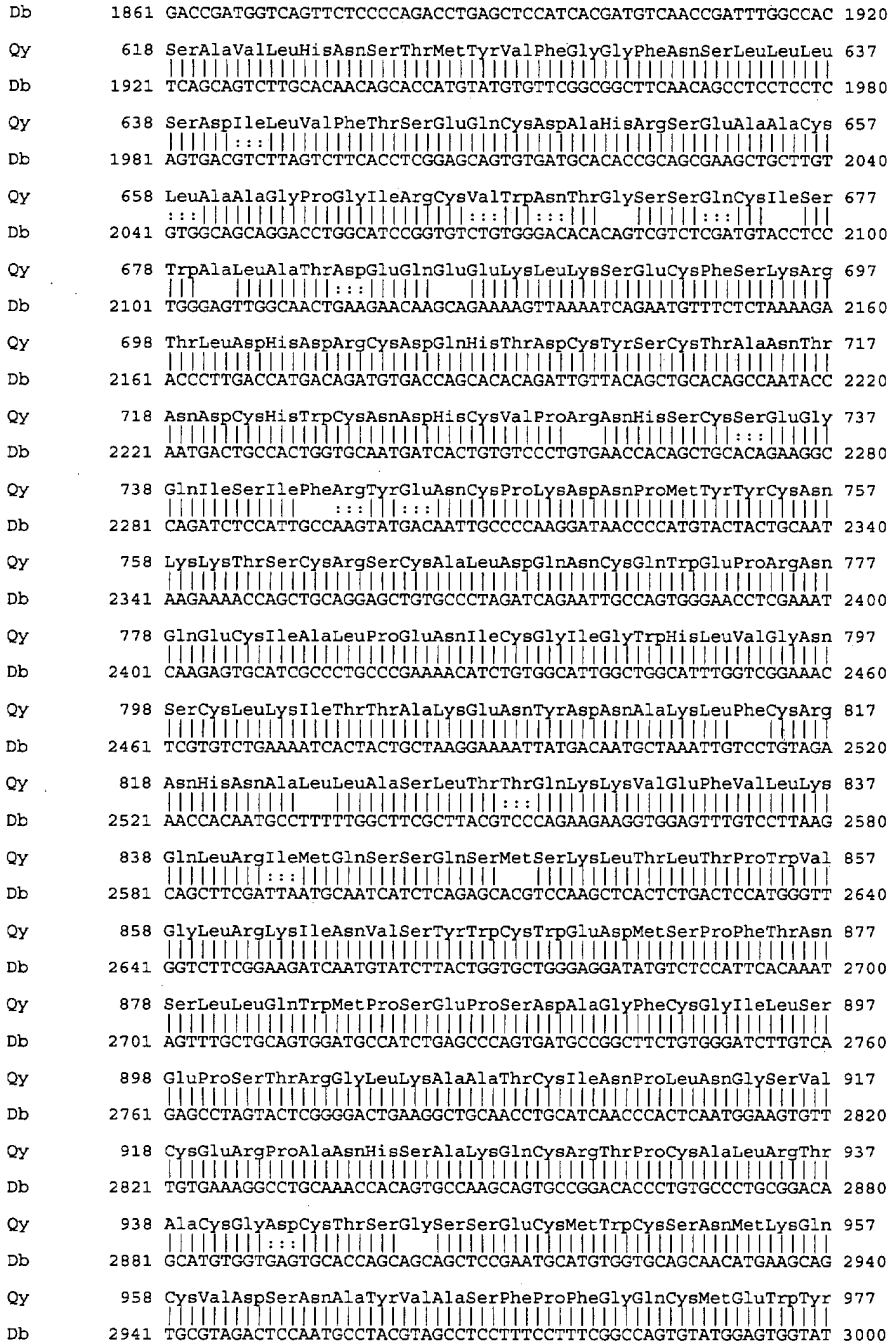
Qy        518 GlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyr 537
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GGTGGCTACAAGGCTTTCAGCGCCAACAAGTACCGGCTTGCAGATGACCTCTACAGATAC 1680

Qy        538 AspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeuHis 557
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CATGTGGACACTCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCAT 1740

Qy        558 ThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThr 577
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 ACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTCGGAGGGAACACACACAATGACACT 1800

Qy        578 SerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAlaCys 597
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 TCCATGAGCCACGGTGCCAAATGCTTCTCCTCAGACTTCATGGCTTATGACATTGCTTGT 1860

Qy        598 AspArgTrpSerValLeuProArgProAspLeuHisHisAspValAsnArgPheGlyHis 617
              ||||||||||||||||||||||||||:::||||||||||||||||||||||||||||||
```


Db 1861 GACCGATGGTCAGTTCTCCCCAGACCTGAGCTCCATCACGATGTCAACCGATTTGGCCAC 1920
Qy 618 SerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeu 637
Db 1921 TCAGCAGTCTTGCACAACAGCACCATGTATGTGTTCGGCGGCTTCAACAGCCTCCTCCTC 1980
Qy 638 SerAspIleLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCys 657
Db 1981 AGTGACGTCTTAGTCTTCACCTCGGAGCAGTGTGATGCACACCGCAGCGAAGCTGCTTGT 2040
Qy 658 LeuAlaAlaGlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIleSer 677
Db 2041 GTGGCAGCAGGACCTGGCATCCGGTGTCTGTGGGACACACAGTCGTCTCGATGTACCTCC 2100
Qy 678 TrpAlaLeuAlaThrAspGluGlnGluGluLysLeuLysSerGluCysPheSerLysArg 697
Db 2101 TGGGAGTTGGCAACTGAAGAACAAGCAGAAAAGTTAAAATCAGAATGTTTCTCTAAAAGA 2160
Qy 698 ThrLeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThr 717
Db 2161 ACCCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAATACC 2220
Qy 718 AsnAspCysHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGluGly 737
Db 2221 AATGACTGCCACTGGTGCAATGATCACTGTGTCCCTGTGAACCACAGCTGCACAGAAGGC 2280
Qy 738 GlnIleSerIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsn 757
Db 2281 CAGATCTCCATTGCCAAGTATGACAATTGCCCCAAGGATAACCCCATGTACTACTGCAAT 2340
Qy 758 LysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsn 777
Db 2341 AAGAAAACCAGCTGCAGGAGCTGTGCCCTAGATCAGAATTGCCAGTGGGAACCTCGAAAT 2400
Qy 778 GlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsn 797
Db 2401 CAAGAGTGCATCGCCCTGCCCGAAAACATCTGTGGCATTGGCTGGCATTTGGTCGGAAAC 2460
Qy 798 SerCysLeuLysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArg 817
Db 2461 TCGTGTCTGAAAATCACTACTGCTAAGGAAAATTATGACAATGCTAAATTGTCCTGTAGA 2520
Qy 818 AsnHisAsnAlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLys 837
Db 2521 AACCACAATGCCTTTTTGGCTTCGCTTACGTCCCAGAAGAAGGTGGAGTTTGTCCTTAAG 2580
Qy 838 GlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpVal 857
Db 2581 CAGCTTCGATTAATGCAATCATCTCAGAGCACGTCCAAGCTCACTCTGACTCCATGGGTT 2640
Qy 858 GlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsn 877
Db 2641 GGTCTTCGGAAGATCAATGTATCTTACTGGTGCTGGGAGGATATGTCTCCATTCACAAAT 2700
Qy 878 SerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSer 897
Db 2701 AGTTTGCTGCAGTGGATGCCATCTGAGCCCAGTGATGCCGGCTTCTGTGGGATCTTGTCA 2760
Qy 898 GluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerVal 917
Db 2761 GAGCCTAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGAAGTGTT 2820
Qy 918 CysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThr 937
Db 2821 TGTGAAAGGCCTGCAAACCACAGTGCCAAGCAGTGCCGGACACCCTGTGCCCTGCGGACA 2880
Qy 938 AlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGln 957
Db 2881 GCATGTGGTGAGTGCACCAGCAGCAGCTCCGAATGCATGTGGTGCAGCAACATGAAGCAG 2940
Qy 958 CysValAspSerAsnAlaTyrValAlaSerProProPheGlyGlnCysMetGluTrpTyr 977
Db 2941 TGCGTAGACTCCAATGCCTACGTAGCCTCCTTTCCTTTCGGCCAGTGTATGGAGTGGTAT 3000
Qy 978 ThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeu 997
Db 3001 ACAATGAGCAGCTGCCCACCTGAAAATTGCTCTGGCTACTGTACCTGCAGCCATTGCTTG 3060
Qy 998 GluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIle 1017
Db 3061 GAGCAGCCAGGCTGTGGTTGGTGTACTGATCCTAGCAATACTGGGAAAGGCAAATGTATT 3120
Qy 1018 GluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyr 1037
Db 3121 GAGGGAAGCTATAAAGGACCCGTGAAGATGCCTTCACATGCCTCTACAGGAAATGTGTAC 3180
Qy 1038 ProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPhe 1057
Db 3181 CCACAGCCCCTTCTGAACTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTT 3240
Qy 1058 IleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCys 1077
Db 3241 ATTCACTGTCCAGCTTGCCAGTGCAATGGACACAGCAAGTGCATCAACCAGAGCATCTGT 3300
Qy 1078 GluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyr 1097
Db 3301 GAGAAGTGTGAGGACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTAT 3360
Qy 1098 GlyAspProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeu 1117
Db 3361 GGTGACCCGACTAATGGAGGCAAATGTCAGCCATGCAAGTGCAACGGGCACGCATCACTG 3420
Qy 1118 CysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCys 1137
Db 3421 TGCAACACCAACACTGGCAAGTGCTTCTGTACCACCAAGGGCGTCAAGGGAGAGGAGTGC 3480
Qy 1138 GlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyr 1157
Db 3481 CAGCTATGTGAGGTAGAAAATCGATACCAGGGAAACCCTCTCAAAGGAACATGTTACTAT 3540
Qy 1158 ThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyr 1177
Db 3541 ACCCTTCTCATTGACTATCAGTTCACTTTTAGCCTGTCCCAGGAAGACGACCGCTATTAC 3600
Qy 1178 ThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIle 1197
Db 3601 ACAGCCATCAATTTTGTGGCTACTCCTGATGAACAAAACAGGGATTTGGACATGTTCATC 3660
Qy 1198 AsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThr 1217
Db 3661 AATGCCTCCAAAAACTTCAACCTCAACATCACCTGGGCCACCAGTTTCCCAGCTGGAACC 3720
Qy 1218 GlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSer 1237
Db 3721 CAGACTGGAGAAGAGGTGCCTGTTGTTTCAAAAACCAACATCAAGGAATACAAAGATAGC 3780
Qy 1238 PheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrVal 1257
Db 3781 TTCTCTAATGAGAAATTTGATTTTCGCAACCATCCAAACATTACTTTCTTTGTTTATGTC 3840
Qy 1258 SerAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMet 1277
Db 3841 AGTAATTTCACTTGGCCCATCAAAATTCAAATTGCCTTCTCCCAGCACAGCAATTTTATG 3900
Qy 1278 AspLeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuLeuValAla 1297
Db 3901 GACCTGGTGCAGTTCTTCGTGACTTTCTTCAGCTGTTTCCTCTCTCTGCTTCTGGTGGCG 3960
Qy 1298 AlaValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeuLeu 1317
Db 3961 GCAGTGGTCTGGAAGATCAAGCAGAGCTGTTGGGCATCCAGGAGGAGAGAGCAACTTCTT 4020
Qy 1318 ArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThr 1337
Db 4021 CGGGAGATGCAACAGATGGCCAGCCGCCCTTTTGCTTCTGTAAATGTTGCCTTGGAAACA 4080

```
Qy       1338 AspGluGluProProAspLeuIleGlyGlySerIleLysThrValProLysProIleAla 1357
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4081 GATGAAGAACCTCCTGATCTCATTGGGGGAAGTATAAAGACTGTTCCTAAGCCCATTGCC 4140

Qy       1358 LeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProArg 1377
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4141 CTGGAGCCCTGCTTTGGTAACAAAGCTGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGA 4200

Qy       1378 GlyLeuGlyGlyIleProProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuVal 1397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4201 GGACTGGGAGGAATCCCTCCTCCTGGTCAGTCAGGTCTTGCTGTGGCCAGTGCCCTGGTG 4260

Qy       1398 AspIleSerGlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAsnArg 1417
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4261 GACATTTCTCAGCAGATGCCAATAGTGTACAAGGAGAAGTCAGGAGCTGTAAGGAACCGG 4320

Qy       1418 LysGlnGlnProProAlaGlnProGlyThrCysIle 1429
             ||||||||||||||||||||||||||||||||||||
Db       4321 AAGCAGCAGCCCCCTGCACAGCCTGGAACCTGCATT 4356
```

RESULT 6
AAZ91916

```
ID   AAZ91916 standard; cDNA; 8827 BP.
XX
AC   AAZ91916;
XX
DT   08-JUN-2000  (first entry)
XX
DE   Wild type (C57BL/6J) murine mahogany protein coding sequence.
XX
KW   Mahogany gene; mouse; mg gene; regulatory defect; gene therapy; obesity;
KW   weight regulation; cell therapy; body weight disorder; cachexia;
KW   anorexia; hyperpigmentation; increased metabolic rate disorder;
KW   hyperphagia; Antiobesity; antianorexic; anticachexic; ss.
XX
OS   Mus sp.
XX
PN   WO200005373-A2.
XX
PD   03-FEB-2000.
XX
PF   21-JUL-1999;   99WO-US016484.
XX
PR   21-JUL-1998;   98US-0093630P.
PR   20-OCT-1998;   98US-0104978P.
PR   05-FEB-1999;   99US-00245041.
XX
PA   (MILL-) MILLENIUM PHARM INC.
XX
PI   Moore K,  Nagle DL;
XX
DR   WPI; 2000-195103/17.
DR   P-PSDB; AAY81803.
XX
PT   New human and murine mahogany genes, useful, e.g. for diagnosis and
PT   treatment of body weight disorders.
XX
PS   Claim 1; Fig 2a; 188pp; English.
XX
CC   This sequence represents a murine mahogany gene of the invention. The
CC   mahogany genes are used: (i) to produce recombinant mahogany (mg)
CC   proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming
CC   therapeutics; (iii) as a source of diagnostic probes and primers for
CC   detecting expression of mg genes or mutations, regulatory defects, in
CC   this gene, or for isolation of related sequences; and (iv) in (cell-
CC   based) gene therapy. (II) are used to raise specific antibodies (Ab); to
CC   identify other (extra)cellular products involved in weight regulation,
CC   and to screen for agents that disrupt interaction between (II) and other
CC   macromolecules. The Ab are used to detect abnormal levels (or function)
CC   of (II) (for diagnosis, prognosis or monitoring of treatment); to
CC   evaluate (II)-expressing cells intended for cell therapy, and as
CC   therapeutic mg inhibitors. Cells that express the mg gene (or contain the
CC   mg polypeptide) are used to identify agents (A) that modulate mg
CC   activity. (A) are potentially useful for the treatment of body weight
CC   disorders, particularly obesity, cachexia or anorexia, or other
CC   conditions associated with the mg gene such as hyperpigmentation,
CC   hyperphagia and disorders that result in increased metabolic rate
XX
SQ   Sequence 8827 BP; 2304 A; 2077 C; 2162 G; 2284 T; 0 U; 0 Other;
```

Alignment Scores:

Pred. No.:	0	Length:	8827
Score:	7294.00	Matches:	1331
Percent Similarity:	90.93%	Conservative:	32
Best Local Similarity:	88.79%	Mismatches:	57
Query Match:	92.15%	Indels:	79
DB:	3	Gaps:	4

US-09-787-097-12 (1-1429) x AAZ91916 (1-8827)

```
Qy          2 ValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAlaAla 21
              ||||||||||||||||||||||||||||||||||||  |||     |||||||||
Db         83 GTGGCGGCGGCGGCGGCGACTGAGGCGCGGCTGAGGGGGAGCACGAGGACGACAGCAGCG 142

Qy         22 LeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArgPro 41
              |||||||||   |||     |||          |||   |||   ||||||
Db        143 CCTGCGGGCAGGAAGGGCAGGCAGCACCGACCCTGCACCGCGACAGGGGCCTGGAGGCCG 202

Qy         42 GlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSer---------------ProPro 56
              |||     |||   |||   |||||||||:::||||||                 ||||||
Db        203 GGACCGCGCGCCCGGCTGTGTCTCCCGCGGGTGCTGTCGCGGGCGCTGCCCCCGCCGCCG 262

Qy         57 LeuArgProArgLeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeu 76
              |||   |||   ||||||       ||||||||||||||||
Db        263 CTGCTGCCG---CTGCTCTTTTCGCTGCTGCTGCTGCCG---------------------- 298

Qy         77 LeuProCysGluAlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAla 96
              ||||||   ||||||||||||||||||   ||||||||||||||||||||||||||||||
Db        299 CTGCCCCGGGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCC 358

Qy         97 GluAlaLysGluCysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGly 116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 GAGGCCAAGGAATGTGACCGGCCGTGTGTCAACGGCGGCCGCTGCAACCCTGGCACCGGC 418

Qy        117 GlnCysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPhe 136
              ||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        419 CAGTGCGTCTGCCCCACGGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTC 478

Qy        136 ------------------------------------------------------------ 136

Db        479 AGGACATCTGTCTCACGCCTATAATCACAGCTGTTCGGAAGGTGAGGCTGGAGGAACAGT 538

Qy        136 ------------------------------------------------------------ 136

Db        539 TCGAGGCAAGCTTCGGCTACAGAATAAGTTCAAGAGTAACCTGGGGCAACTTGGGCTTGT 598

Qy        136 ------------------------------------------------------------ 136

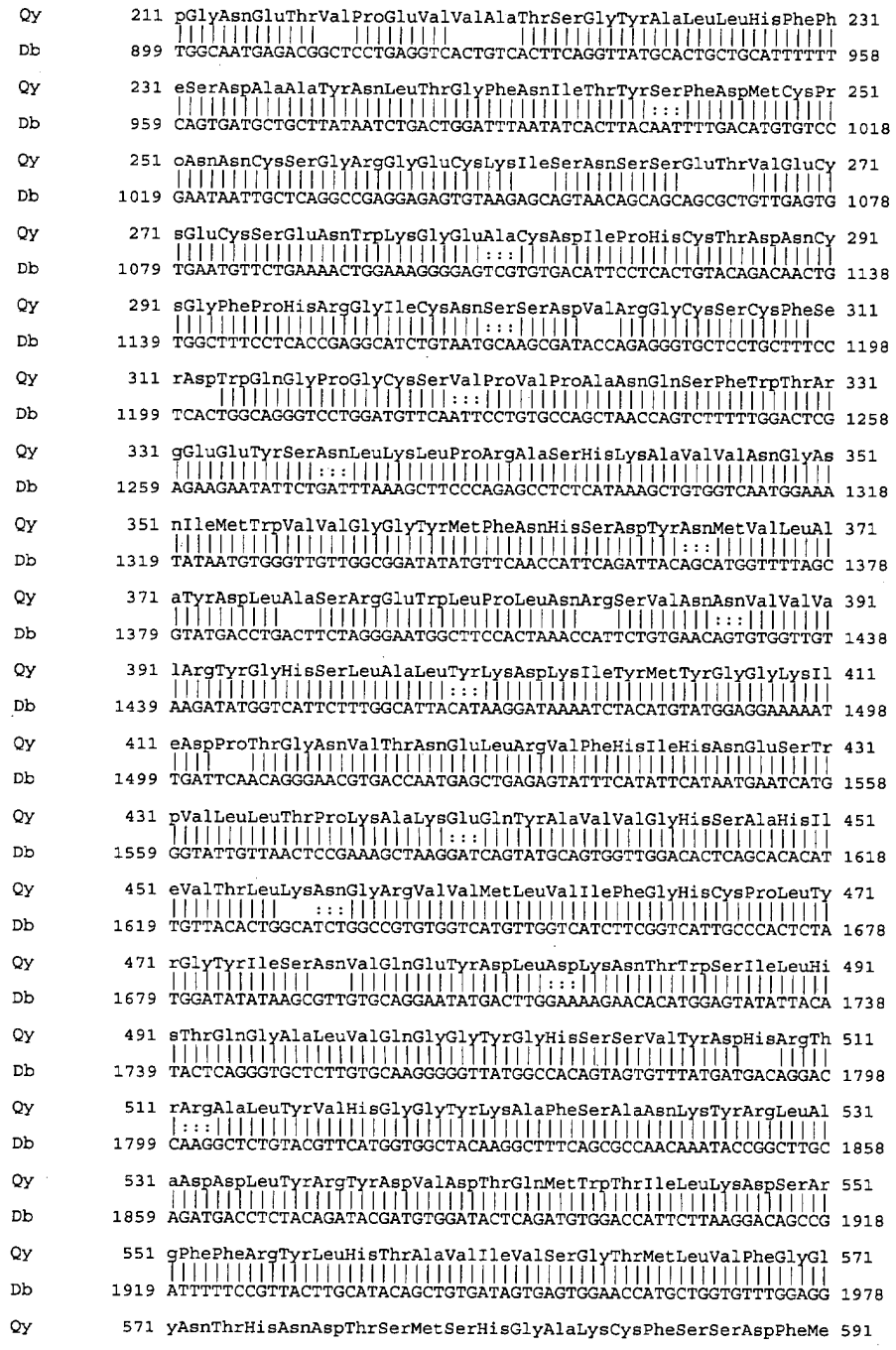
Db        599 CTCCAAAACCAAAATGAGCGAAAAGGAGCAAGCTAGAGTCTTTTGGGAAAATTTTAGCTG 658

Qy        137 --------------Arg-LeuThrGlySerSerGlyPheValThrAspGlyProGlyAs 151
                           ||| ||||||||||||||||||||||||||||||||||||||||||
Db        659 ACTAATTTTTCACCGAGAACTAACTGGCTCTTCTGGATTTGTAACAGATGGACCTGGGAA 718

Qy        151 nTyrLysTyrLysThrLysCysThrTrpLeuIleGluGlyGlnProAsnArgIleMetAr 171
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 TTATAAATATAAGACGAAGTGCACATGGCTCATTGAAGGACAGCCAAATAGAATAATGAG 778

Qy        171 gLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGl 191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 ACTTCGCTTCAACCATTTTGCTACAGAATGTAGCTGGGACCATTTATATGTTTATGATGG 838

Qy        191 yAspSerIleTyrAlaProLeuValAlaAlaPheSerGlyLeuIleValProGluArgAs 211
              ||||||||||||||||||||:::||||||||||||||||||||||||||||||||||||||
Db        839 GGACTCAATCTACGCACCTCTGATTGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGA 898
```


Qy 211 pGlyAsnGluThrValProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePh 231
Db 899 TGGCAATGAGACGGCTCCTGAGGTCACTGTCACTTCAGGTTATGCACTGCTGCATTTTTT 958
Qy 231 eSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysPr 251
Db 959 CAGTGATGCTGCTTATAATCTGACTGGATTTAATATCACTTACAATTTTGACATGTGTCC 1018
Qy 251 oAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSerSerGluThrValGluCy 271
Db 1019 GAATAATTGCTCAGGCCGAGGAGAGTGTAAGAGCAGTAACAGCAGCAGCGCTGTTGAGTG 1078
Qy 271 sGluCysSerGluAsnTrpLysGlyGluAlaCysAspIleProHisCysThrAspAsnCy 291
Db 1079 TGAATGTTCTGAAAACTGGAAAGGGGAGTCGTGTGACATTCCTCACTGTACAGACAACTG 1138
Qy 291 sGlyPheProHisArgGlyIleCysAsnSerSerAspValArgGlyCysSerCysPheSe 311
Db 1139 TGGCTTTCCTCACCGAGGCATCTGTAATGCAAGCGATACCAGAGGGTGCTCCTGCTTTCC 1198
Qy 311 rAspTrpGlnGlyProGlyCysSerValProValProAlaAsnGlnSerPheTrpThrAr 331
Db 1199 TCACTGGCAGGGTCCTGGATGTTCAATTCCTGTGCCAGCTAACCAGTCTTTTTGGACTCG 1258
Qy 331 gGluGluTyrSerAsnLeuLysLeuProArgAlaSerHisLysAlaValValAsnGlyAs 351
Db 1259 AGAAGAATATTCTGATTTAAAGCTTCCCAGAGCCTCTCATAAAGCTGTGGTCAATGGAAA 1318
Qy 351 nIleMetTrpValValGlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAl 371
Db 1319 TATAATGTGGGTTGTTGGCGGATATATGTTCAACCATTCAGATTACAGCATGGTTTTAGC 1378
Qy 371 aTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSerValAsnAsnValValVa 391
Db 1379 GTATGACCTGACTTCTAGGGAATGGCTTCCACTAAACCATTCTGTGAACAGTGTGGTTGT 1438
Qy 391 lArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLysIl 411
Db 1439 AAGATATGGTCATTCTTTGGCATTACATAAGGATAAAATCTACATGTATGGAGGAAAAAT 1498
Qy 411 eAspProThrGlyAsnValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTr 431
Db 1499 TGATTCAACAGGGAACGTGACCAATGAGCTGAGAGTATTTCATATTCATAATGAATCATG 1558
Qy 431 pValLeuLeuThrProLysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIl 451
Db 1559 GGTATTGTTAACTCCGAAAGCTAAGGATCAGTATGCAGTGGTTGGACACTCAGCACACAT 1618
Qy 451 eValThrLeuLysAsnGlyArgValValMetLeuValIlePheGlyHisCysProLeuTy 471
Db 1619 TGTTACACTGGCATCTGGCCGTGTGGTCATGTTGGTCATCTTCGGTCATTGCCCACTCTA 1678
Qy 471 rGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHi 491
Db 1679 TGGATATATAAGCGTTGTGCAGGAATATGACTTGGAAAAGAACACATGGAGTATATTACA 1738
Qy 491 sThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisSerServalTyrAspHisArgTh 511
Db 1739 TACTCAGGGTGCTCTTGTGCAAGGGGGTTATGGCCACAGTAGTGTTTATGATGACAGGAC 1798
Qy 511 rArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAl 531
Db 1799 CAAGGCTCTGTACGTTCATGGTGGCTACAAGGCTTTCAGCGCCAACAAATACCGGCTTGC 1858
Qy 531 aAspAspLeuTyrArgTyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerAr 551
Db 1859 AGATGACCTCTACAGATACGATGTGGATACTCAGATGTGGACCATTCTTAAGGACAGCCG 1918
Qy 551 gPhePheArgTyrLeuHisThrAlaValIleValSerGlyThrMetLeuValPheGlyGl 571
Db 1919 ATTTTTCCGTTACTTGCATACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGAGG 1978
Qy 571 yAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMe 591
Db 1979 GAACACACACAATGACACTTCCATGAGCCACGGTGCCAAATGCTTCTCCTCGGACTTCAT 2038
Qy 591 tAlaTyrAspIleAlaCysAspArgTrpSerValLeuProArgProAspLeuHisHisAs 611
Db 2039 GGCTTATGACATTGCTTGTGACCGATGGTCAGTGCTTCCCAGACCTGAGCTCCATCATGA 2098
Qy 611 pValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGl 631
Db 2099 TGTCAACAGATTTGGCCATTCAGCAGTCTTGTACAACAGCACCATGTATGTGTTCGGCGG 2158
Qy 631 yPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHi 651
Db 2159 CTTCAACAGCCTCCTCCTCAGTGACGTCTTGGTCTTTACCTCGGAGCAGTGCGATGCACA 2218
Qy 651 sArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgCysValTrpAsnThrGl 671
Db 2219 CCGCAGTGAAGCTGCTTGTGTGGCAGCAGGACCTGGTATCCGGTGTCTGTGGGACACACA 2278
Qy 671 ySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnGluGluLysLeuLysSe 691
Db 2279 GTCGTCTCGATGTACCTCCTGGGAGTTGGCAACTGAAGAACAAGCAGAAAAGTTAAAATC 2338
Qy 691 rGluCysPheSerLysArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTy 711
Db 2339 AGAGTGTTTTTCTAAAAGAACCCTTGACCATGACAGATGTGACCAGCACACAGATTGTTA 2398
Qy 711 rSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAspHisCysValProArgAs 731
Db 2399 CAGCTGCACAGCCAATACCAATGACTGCCACTGGTGCAATGATCACTGTGTCCCTGTGAA 2458
Qy 731 nHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGluAsnCysProLysAspAs 751
Db 2459 CCACAGCTGCACAGAAGGCCAGATCTCCATTGCCAAGTATGAGAGTTGCCCCAAGGATAA 2518
Qy 751 nProMetTyrTyrCysAsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCy 771
Db 2519 CCCCATGTACTACTGCAATAAGAAAACCAGCTGCAGGAGCTGTGCCCTAGACCAGAACTG 2578
Qy 771 sGlnTrpGluProArgAsnGlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGl 791
Db 2579 CCAGTGGGAGCCCCGGAATCAAGAGTGCATCGCCCTGCCGGAAAATATCTGTGGCAATGG 2638
Qy 791 yTrpHisLeuValGlyAsnSerCysLeuLysIleThrThrAlaLysGluAsnTyrAspAs 811
Db 2639 CTGGCATTTGGTTGGAAACTCGTGTCTGAAAATCACTACTGCTAAGGAGAATTATGACAA 2698
Qy 811 nAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSerLeuThrThrGlnLysLy 831
Db 2699 TGCTAAATTGTCCTGTAGGAACCACAATGCCTTTTTGGCTTCCCTCACATCCCAGAAGAA 2758
Qy 831 sValGluPheValLeuLysGlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLe 851
Db 2759 GGTGGAGTTTGTCCTTAAGCAGCTTCGATTAATGCAATCATCTCAAAGTATGTCCAAGCT 2818
Qy 851 uThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAs 871
Db 2819 CACTCTGACTCCATGGGTTGGTCTTCGGAAGATCAATGTGTCTTACTGGTGCTGGGAGGA 2878
Qy 871 pMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSerAspAlaGl 891
Db 2879 TATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGTGATGCTGG 2938
Qy 891 yPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThrCysIleAs 911
Db 2939 CTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGATTAAAGGCTGCAACCTGCATCAA 2998
Qy 911 nProLeuAsnGlySerValCysGluArgProAlaAsnHisSerAlaLysGlnCysArgTh 931
Db 2999 CCCTCTCAATGGCAGCGTCTGTGAAAGGCCTGCAAACCACAGTGCCAAGCAGTGCCGGAC 3058
Qy 931 rProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTr 951

```
Db       3059 ACCATGTGCCCTGCGGACAGCGTGTGGCGAGTGCACTAGCAGCAGCTCGGAGTGCATGTG 3118
Qy        951 pCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrValAlaSerPheProPheGl 971
Db       3119 GTGCAGTAACATGAAGCAGTGTGTGGACTCCAATGCCTACGTGGCCTCCTTCCCTTTTGG 3178
Qy        971 yGlnCysMetGluTrpTyrThrMetSerThrCysProProGluAsnCysSerGlyTyrCy 991
Db       3179 CCAGTGTATGGAATGGTATACGATGAGCAGCTGCCCACCTGAAAATTGCTCTGGCTACTG 3238
Qy        991 sThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCysThrAspProSerAsnTh 1011
Db       3239 TACCTGCAGCCATTGCTTGGAGCAGCCAGGCTGTGGTTGGTGTACTGATCCTAGCAATAC 3298
Qy       1011 rGlyLysGlyLysCysIleGluGlySerTyrLysGlyProValLysMetProSerGlnAl 1031
Db       3299 TGGGAAAGGAAAATGTATTGAGGGCAGCTATAAAGGACCTGTGAAGATGCCGTCACAGGC 3358
Qy       1031 aProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSe 1051
Db       3359 CTCTGCAGGAAATGTGTATCCACAGCCCCTTCTGAACTCCAGCATGTGTCTAGAGGACAG 3418
Qy       1051 rArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCy 1071
Db       3419 CAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAGTGCAACGGACACAGCAAATG 3478
Qy       1071 sIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrThrGlyLysHisCysGluTh 1091
Db       3479 CATCAACCAGAGTATCTGTGAGAAGTGTGAGGACCTGACCACGGGCAAGCACTGCGAGAC 3538
Qy       1091 rCysIleSerGlyPheTyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLysCy 1111
Db       3539 CTGCATATCTGGCTTCTATGGTGACCCGACTAATGGAGGCAAATGTCAGCCATGCAAGTG 3598
Qy       1111 sAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGl 1131
Db       3599 CAATGGGCACGCATCACTGTGCAACACCAACACCGGCAAGTGCTTCTGTACCACCAAAGG 3658
Qy       1131 yValLysGlyAspGluCysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLe 1151
Db       3659 TGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCT 3718
Qy       1151 uArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGl 1171
Db       3719 CAAAGGAACATGCTACTATACCCTTCTCATTGACTATCAGTTCACCTTTAGCCTGTCCCA 3778
Qy       1171 nGluAspAspArgTyrTyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnAr 1191
Db       3779 GGAAGACGACCGCTACTACACAGCCATCAACTTTGTGGCTACTCCTGATGAACAAAACAG 3838
Qy       1191 gAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAl 1211
Db       3839 GGATTTGGACATGTTCATCAATGCCTCCAAAAACTTCAACCTCAACATCACCTGGGCCAC 3898
Qy       1211 aSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIl 1231
Db       3899 CAGCTTCCCAGCCGGAACCCAGACTGGAGAAGAGGTGCCTGTTGTTTCAAAAACCAACAT 3958
Qy       1231 eLysGluTyrLysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIl 1251
Db       3959 CAAGGAATACAAAGATAGCTTCTCTAATGAGAAATTTGATTTTCGCAACGATCCAAACAT 4018
Qy       1251 eThrPhePheValTyrValSerAsnPheThrTrpProIleLysIleGlnIleAlaPheSe 1271
Db       4019 CACTTTCTTTGTTTATGTCAGTAATTTCACTTGGCCCATCAAAATTCAGATTGCCTTCTC 4078
Qy       1271 rGlnHisSerAsnPheMetAspLeuValGlnPhePheValThrPhePheSerCysPheLe 1291
Db       4079 CCAGCACAGCAACTTCATGGACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCT 4138
Qy       1291 uSerLeuLeuLeuValAlaAlaValValTrpLysIleLysGlnSerCysTrpAlaSerAr 1311
Db       4139 CTCGCTGCTTCTGGTGGCTGCAGTGGTCTGGAAGATCAAGCAGAGCTGTTGGGCATCCAG 4198
```

```
Qy       1311 gArgArgGluGlnLeuLeuArgGluMetGlnGlnMetAlaSerArgProPheAlaSerVa 1331
Db       4199 GCGGAGAGAGCAACTTCTTCGGGAGATGCAACAGATGGCCAGCCGCCCCTTTGCTTCTGT 4258
Qy       1331 lAsnValAlaLeuGluThrAspGluGluProProAspLeuIleGlyGlySerIleLysTh 1351
Db       4259 AAACGTTGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGAC 4318
Qy       1351 rValProLysProIleAlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerVa 1371
Db       4319 TGTTCCCAAACCCATTGCACTGGAGCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGT 4378
Qy       1371 lPheValArgLeuProArgGlyLeuGlyGlyIleProProProGlyGlnSerGlyLeuAl 1391
Db       4379 GTTTGTGAGGCTCCCTCGAGGCCTGGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGC 4438
Qy       1391 aValAlaSerAlaLeuValAspIleSerGlnGlnMetProIleValTyrLysGluLysSe 1411
Db       4439 TGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTC 4498
Qy       1411 rGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnProGlyThrCysIle 1429
Db       4499 AGGAGCCGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGCCTGGGACCTGCATC 4553
```

RESULT 7
AAZ52277
ID AAZ52277 standard; cDNA; 3819 BP.
XX
AC AAZ52277;
XX
DT 18-JUL-2000 (first entry)
XX
DE Human soluble attractin-2 cDNA.
XX
KW Human; soluble attractin-2; immune response; macrophage; monocyte;
KW T cell; immunostimulant; immunosuppressed patient; cancer;
KW immunodeficiency syndrome; transplant; autoimmune disease; ss.
XX
OS Homo sapiens.
XX
FH Key   Location/Qualifiers
FT CDS   1. .3819
FT   /*tag= a
FT   /product= "Soluble attractin-2"
XX
PN WO200015651-A1.
XX
PD 23-MAR-2000.
XX
PF 14-SEP-1999; 99WO-US020948.
XX
PR 14-SEP-1998; 98US-0100137P.
XX
PA (DAND) DANA FARBER CANCER INST INC.
XX
PI Duke-Cohan JS, Schlossman SF;
XX
DR WPI; 2000-271373/23.
DR P-PSDB; AAY70692.
XX
PT Isolated nucleic acids encoding human attractin polypeptides useful for
PT enhancing immune responses.
XX
PS Claim 3; Fig 12; 120pp; English.
XX
CC The patent discloses four forms of human attractin polypeptides which
CC enhance immune response by promoting macrophage and monocyte spreading in
CC the presence of T cells. These include soluble attractin-1 and -2 and
CC membrane attractin-1 and -2. These various forms of attractin are encoded
CC by alternatively spliced mRNA molecule transcribed from a single gene.
CC The present sequence is a cDNA encoding human soluble attractin-2.
CC Attractin-2 differs from attractin-1 in having a 74 amino acid insert in

CC the N-terminal. This sequence can be used to enhance immune response in
CC immunosuppressed patients such as those undergoing chemo- and radio-
CC therapy treatment for cancer or those suffering from common variable
CC immunodeficiency syndrome. The proteins may also be used to screen
CC modulators (agonists and antagonists) of immune response which may also
CC be used to regulate immune reactions. Attractin antibodies can be used to
CC inhibit immune response in transplant recipients or patients afflicted
CC with autoimmune disease
XX
SQ Sequence 3819 BP; 969 A; 899 C; 991 G; 960 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	3819
Score:	7089.00	Matches:	1268
Percent Similarity:	99.76%	Conservative:	1
Best Local Similarity:	99.69%	Mismatches:	3
Query Match:	89.56%	Indels:	0
DB:	3	Gaps:	0

US-09-787-097-12 (1-1429) x AAZ52277 (1-3819)

```
Qy          1 MetValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 60

Qy         21 AlaLeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArg 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGGGACTGGGACGTGACCAGGGCTGGGAGG 120

Qy         41 ProGlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSerProProLeuArgProArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGGGGCTGGGGGCCGGGCTGCGCCTCCCGCGGCTGCTGTCTCCACCGCTGCGGCCACGG 180

Qy         61 LeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeuProCysGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGCTGCTGCTGCTGTTGTTGCTCCCGCCGCCGCTGTTGCTGCTGCTGCTGCCCTGTGAG 240

Qy         81 AlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCGAGGCCGCGGCGGCGGCGGCGGCGGTGTCGGGCTCAGCCGCAGCCGAGGCCAAGGAA 300

Qy        101 CysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCys 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGTGACCGGCCCTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGC 360

Qy        121 ProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGly 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGA 420

Qy        141 SerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrp 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGG 480

Qy        161 LeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAG 540

Qy        181 CysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAla 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCT 600

Qy        201 AlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGluValVal 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTT 660

Qy        221 AlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGly 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGA 720

Qy        241 PheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCys 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGT 780

Qy        261 LysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsnTrpLysGlyGlu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAGATCAGTAATAGCAGCGAAACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAA 840

Qy        281 AlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsn 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAAT 900

Qy        301 SerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerVal 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTT 960

Qy        321 ProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuPro 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCC 1020

Qy        341 ArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMet 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATG 1080

Qy        361 PheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeu 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTT 1140

Qy        381 ProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATAC 1200

Qy        401 LysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsnValThrAsnGlu 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AAGGATAAAATTTACATGTATGGAGGAAAAATTGATCCAACTGGGAATGTGACCAATGAG 1260

Qy        421 LeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGlu 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGGAG 1320

Qy        441 GlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValVal 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTC 1380

Qy        461 MetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyr 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATAT 1440

Qy        481 AspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGly 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGT 1500

Qy        501 TyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyr 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTAC 1560

Qy        521 LysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAsp 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGAT 1620

Qy        541 ThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaVal 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 ACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTG 1680

Qy        561 IleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSer 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 ATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGC 1740

Qy        581 HisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrp 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGG 1800

Qy        601 SerValLeuProArgProAspLeuHisHisAspValAsnArgPheGlyHisSerAlaVal 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 TCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTC 1860

Qy        621 LeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIle 640
```

```
Db       1861 TTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATC 1920
Qy        641 LeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAla 660
Db       1921 CTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCA 1980
Qy        661 GlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeu 680
Db       1981 GGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTG 2040
Qy        681 AlaThrAspGluGlnGluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAsp 700
Db       2041 GCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGAC 2100
Qy        701 HisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCys 720
Db       2101 CATGACAGATGTGACCAGCACACAGATTGTTACAGCTGTACAGCCAACACCAATGACTGC 2160
Qy        721 HisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSer 740
Db       2161 CACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCC 2220
Qy        741 IlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThr 760
Db       2221 ATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCCATGTACTACTGTAACAAGAAGACC 2280
Qy        761 SerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCys 780
Db       2281 AGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGC 2340
Qy        781 IleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeu 800
Db       2341 ATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTG 2400
Qy        801 LysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsn 820
Db       2401 AAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAAT 2460
Qy        821 AlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArg 840
Db       2461 GCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGA 2520
Qy        841 IleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArg 860
Db       2521 ATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGG 2580
Qy        861 LysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeu 880
Db       2581 AAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTA 2640
Qy        881 GlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSer 900
Db       2641 CAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGT 2700
Qy        901 ThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArg 920
Db       2701 ACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGG 2760
Qy        921 ProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGly 940
Db       2761 CCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGA 2820
Qy        941 AspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAsp 960
Db       2821 GATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGAC 2880
Qy        961 SerAsnAlaTyrValAlaSerProProPheGlyGlnCysMetGluTrpTyrThrMetSer 980
Db       2881 TCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGC 2940
Qy        981 ThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPro 1000
```

```
Db       2941 ACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCA 3000
Qy        1001 GlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySer 1020
Db       3001 GGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCC 3060
Qy        1021 TyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnPro 1040
Db       3061 TATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCC 3120
Qy        1041 LeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCys 1060
Db       3121 CTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGT 3180
Qy        1061 ProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCys 1080
Db       3181 CCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGT 3240
Qy        1081 GluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspPro 1100
Db       3241 GAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCC 3300
Qy        1101 ThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThr 1120
Db       3301 ACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACC 3360
Qy        1121 AsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCys 1140
Db       3361 AACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGT 3420
Qy        1141 GluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeu 1160
Db       3421 GAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTT 3480
Qy        1161 IleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIle 1180
Db       3481 ATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATC 3540
Qy        1181 AsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSer 1200
Db       3541 AATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCC 3600
Qy        1201 LysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGly 1220
Db       3601 AAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGA 3660
Qy        1221 GluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsn 1240
Db       3661 GAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAAT 3720
Qy        1241 GluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPhe 1260
Db       3721 GAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTC 3780
Qy        1261 ThrTrpProIleLysIleGlnIleAlaPheSerGln 1272
Db       3781 ACCTGGCCCATCAAAATTCAGGTGCAAACTGAACAA 3816
```

RESULT 8
AAZ91921
ID AAZ91921 standard; cDNA; 4072 BP.
XX
AC AAZ91921;
XX
DT 08-JUN-2000 (first entry)
XX
DE Human mahogany protein coding sequence #3.
XX
KW Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;
KW weight regulation; cell therapy; body weight disorder; cachexia;
KW anorexia; hyperpigmentation; increased metabolic rate disorder;
KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.
XX

OS Homo sapiens.
XX
PN WO200005373-A2.
XX
PD 03-FEB-2000.
XX
PF 21-JUL-1999; 99WO-US016484.
XX
PR 21-JUL-1998; 98US-0093630P.
PR 20-OCT-1998; 98US-0104978P.
PR 05-FEB-1999; 99US-00245041.
XX
PA (MILL-) MILLENIUM PHARM INC.
XX
PI Moore K, Nagle DL;
XX
DR WPI; 2000-195103/17.
DR P-PSDB; AAY81808.
XX
PT New human and murine mahogany genes, useful, e.g. for diagnosis and
PT treatment of body weight disorders.
XX
PS Claim 1; Fig 19a; 188pp; English.
XX
CC This sequence represents a human mahogany gene of the invention. The
CC mahogany genes are used: (i) to produce recombinant mahogany (mg)
CC proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming
CC therapeutics; (iii) as a source of diagnostic probes and primers for
CC detecting expression of mg genes or mutations, regulatory defects, in
CC this gene, or for isolation of related sequences; and (iv) in (cell-
CC based) gene therapy. (II) are used to raise specific antibodies (Ab); to
CC identify other (extra)cellular products involved in weight regulation,
CC and to screen for agents that disrupt interaction between (II) and other
CC macromolecules. The Ab are used to detect abnormal levels (or function)
CC of (II) (for diagnosis, prognosis or monitoring of treatment); to
CC evaluate (II)-expressing cells intended for cell therapy, and as
CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the
CC mg polypeptide) are used to identify agents (A) that modulate mg
CC activity. (A) are potentially useful for the treatment of body weight
CC disorders, particularly obesity, cachexia or anorexia, or other
CC conditions associated with the mg gene such as hyperpigmentation,
CC hyperphagia and disorders that result in increased metabolic rate
XX
SQ Sequence 4072 BP; 1087 A; 940 C; 1045 G; 1000 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	4072
Score:	6870.50	Matches:	1240
Percent Similarity:	97.41%	Conservative:	3
Best Local Similarity:	97.18%	Mismatches:	20
Query Match:	86.80%	Indels:	13
DB:	3	Gaps:	3

US-09-787-097-12 (1-1429) x AAZ91921 (1-4072)

```
Qy          2 ValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAlaAla 21
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         10 GTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCG 69

Qy         22 LeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArgPro 41
              |||||||||||||||||||||||||||||                |||   |||   ||||||
Db         70 CTCGCGGGCAGGAGCGGCGGGCCGCACCGACCCTGCACCGCGACAGGGGCCTGGAGGCCG 129

Qy         42 GlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSer----------------ProPro 56
              |||     |||   |||   |||||||||:::||||||                ||||||
Db        130 GGACCGCGCGCCCCGGCTGTGTCTCCCGCGGGTGCTGTCGCGGGCGCTGCCCCCGCCGCCG 189

Qy         57 LeuArgProArgLeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeu 76
              |||   |||   ||||||      ||||||||||||||||
Db        190 CTGCTGCCG---CTGCTCTTTTCGCTGCTGCTGCTGCCG---------------------- 225

Qy         77 LeuProCysGluAlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAla 96
              ||||||   ||||||||||||||||||   ||||||||||||||||||||||||||||||
```

```
Db        226 CTGCCCCGGGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCC 285

Qy         97 GluAlaLysGluCysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGly 116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 GAGGCCAAGGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGC 345

Qy        117 GlnCysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPhe 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 CAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTC 405

Qy        137 ArgLeuThrGlySerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThr 156
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        406 AGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACG 465

Qy        157 LysCysThrTrpLeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHis 176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        466 AAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCAT 525

Qy        177 PheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAla 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        526 TTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCA 585

Qy        197 ProLeuValAlaAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrVal 216
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        586 CCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTC 645

Qy        217 ProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyr 236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        646 CCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTAT 705

Qy        237 AsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGly 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        706 AATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGC 765

Qy        257 ArgGlyGluCysLysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsn 276
              ||||||||||||||||||||||||||||||||||||:::|||||||||||||||||||||||
Db        766 CGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAAC 825

Qy        277 TrpLysGlyGluAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArg 296
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        826 TGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGA 885

Qy        297 GlyIleCysAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyPro 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        886 GGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCT 945

Qy        317 GlyCysSerValProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsn 336
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        946 GGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAAC 1005

Qy        337 LeuLysLeuProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValVal 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1006 TTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTT 1065

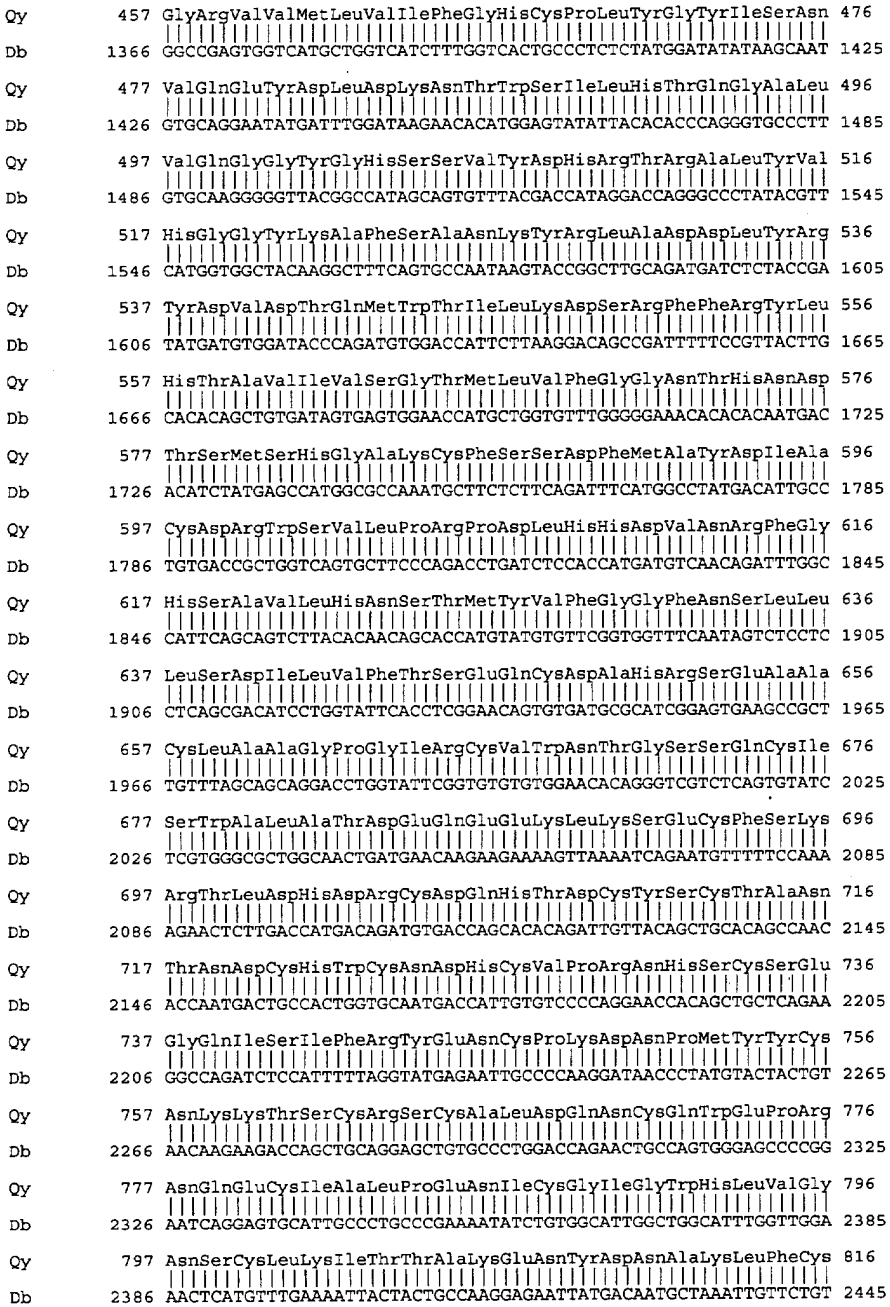
Qy        357 GlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSer 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1066 GGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCT 1125

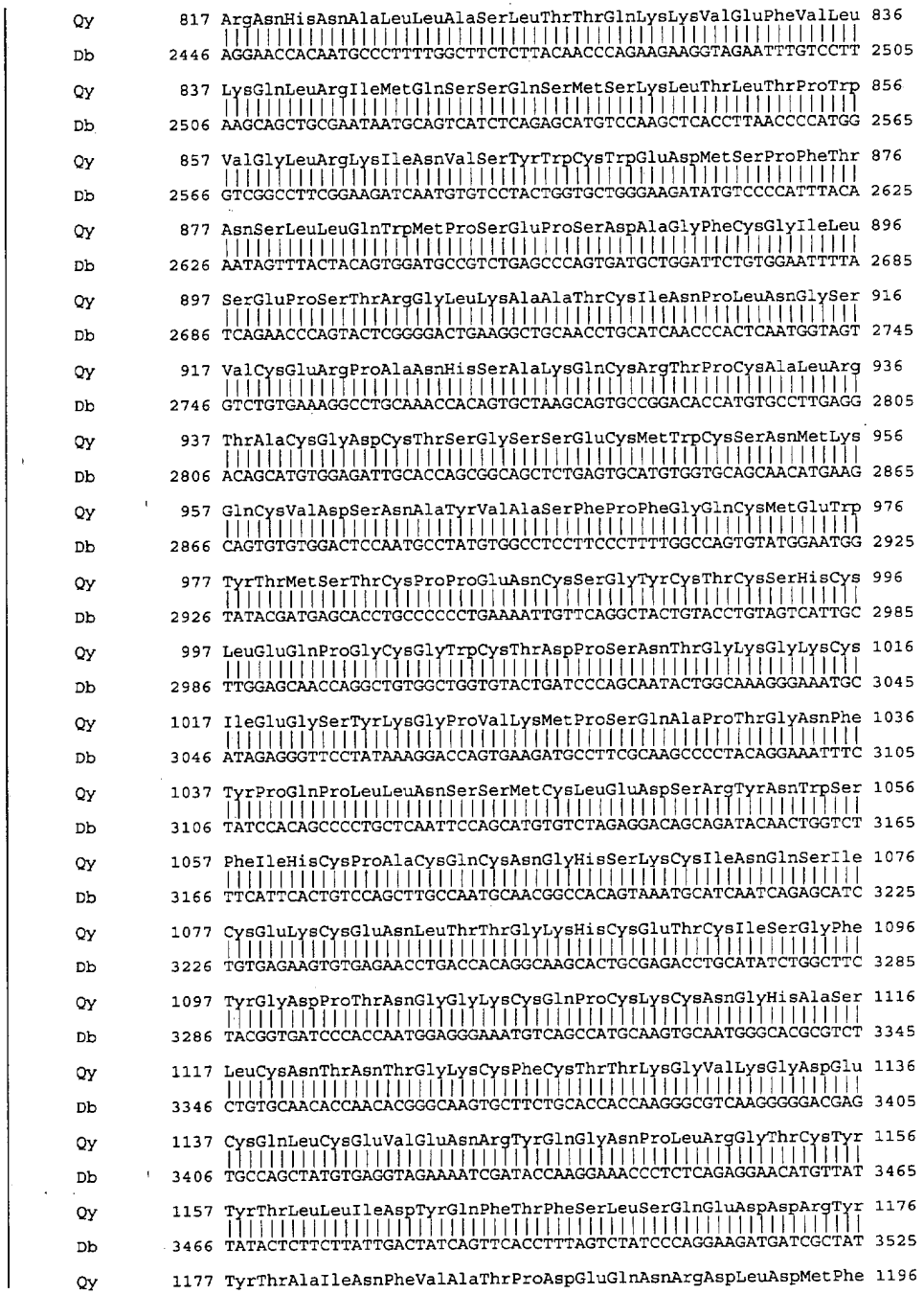
Qy        377 ArgGluTrpLeuProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSer 396
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1126 AGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCT 1185

Qy        397 LeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsn 416
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1186 TTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAAT 1245

Qy        417 ValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrPro 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1246 GTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCT 1305

Qy        437 LysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsn 456
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1306 AAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAAT 1365
```


Qy 457 GlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsn 476
Db 1366 GGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAAT 1425
Qy 477 ValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeu 496
Db 1426 GTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTT 1485
Qy 497 ValGlnGlyGlyTyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrVal 516
Db 1486 GTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTT 1545
Qy 517 HisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArg 536
Db 1546 CATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGA 1605
Qy 537 TyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeu 556
Db 1606 TATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTCCGTTACTTG 1665
Qy 557 HisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAsp 576
Db 1666 CACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGAC 1725
Qy 577 ThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAla 596
Db 1726 ACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCC 1785
Qy 597 CysAspArgTrpSerValLeuProArgProAspLeuHisHisAspValAsnArgPheGly 616
Db 1786 TGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGC 1845
Qy 617 HisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeu 636
Db 1846 CATTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTC 1905
Qy 637 LeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAla 656
Db 1906 CTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCT 1965
Qy 657 CysLeuAlaAlaGlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIle 676
Db 1966 TGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATC 2025
Qy 677 SerTrpAlaLeuAlaThrAspGluGlnGluGluLysLeuLysSerGluCysPheSerLys 696
Db 2026 TCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAA 2085
Qy 697 ArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsn 716
Db 2086 AGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAAC 2145
Qy 717 ThrAsnAspCysHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGlu 736
Db 2146 ACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAA 2205
Qy 737 GlyGlnIleSerIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCys 756
Db 2206 GGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGT 2265
Qy 757 AsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArg 776
Db 2266 AACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGG 2325
Qy 777 AsnGlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGly 796
Db 2326 AATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGA 2385
Qy 797 AsnSerCysLeuLysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCys 816
Db 2386 AACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGT 2445


Qy 817 ArgAsnHisAsnAlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeu 836
Db 2446 AGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTT 2505
Qy 837 LysGlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrp 856
Db 2506 AAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGG 2565
Qy 857 ValGlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThr 876
Db 2566 GTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACA 2625
Qy 877 AsnSerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeu 896
Db 2626 AATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTA 2685
Qy 897 SerGluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySer 916
Db 2686 TCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGT 2745
Qy 917 ValCysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArg 936
Db 2746 GTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGG 2805
Qy 937 ThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLys 956
Db 2806 ACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAG 2865
Qy 957 GlnCysValAspSerAsnAlaTyrValAlaSerProProGlyGlnCysMetGluTrp 976
Db 2866 CAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGG 2925
Qy 977 TyrThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCys 996
Db 2926 TATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGC 2985
Qy 997 LeuGluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCys 1016
Db 2986 TTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGC 3045
Qy 1017 IleGluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPhe 1036
Db 3046 ATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTC 3105
Qy 1037 TyrProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSer 1056
Db 3106 TATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCT 3165
Qy 1057 PheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIle 1076
Db 3166 TTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATC 3225
Qy 1077 CysGluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPhe 1096
Db 3226 TGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTC 3285
Qy 1097 TyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSer 1116
Db 3286 TACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCT 3345
Qy 1117 LeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGlu 1136
Db 3346 CTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAG 3405
Qy 1137 CysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyr 1156
Db 3406 TGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTAT 3465
Qy 1157 TyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyr 1176
Db 3466 TATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTAT 3525
Qy 1177 TyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPhe 1196

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3526 TACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTC 3585

Qy       1197 IleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGly 1216
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3586 ATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGA 3645

Qy       1217 ThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAsp 1236
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3646 ACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGAT 3705

Qy       1237 SerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyr 1256
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3706 AGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTAT 3765

Qy       1257 ValSerAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGln 1272
                ||||||||||||||||||||||||||||||||:::         |||
Db       3766 GTCAGTAATTTCACCTGGCCCATCAAAATTCAGGTGCAAACTGAACAA 3813

RESULT 9
AAZ52274
ID   AAZ52274 standard; cDNA; 3597 BP.
XX
AC   AAZ52274;
XX
DT   18-JUL-2000  (first entry)
XX
DE   Human soluble attractin-1 cDNA.
XX
KW   Human; soluble attractin-1; immune response; macrophage; monocyte;
KW   T cell; immunostimulant; immunosuppressed patient; cancer;
KW   immunodeficiency syndrome; transplant; autoimmune disease; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .3597
FT                   /*tag=  a
FT                   /product= "Soluble attractin-1"
XX
PN   WO200015651-A1.
XX
PD   23-MAR-2000.
XX
PF   14-SEP-1999;   99WO-US020948.
XX
PR   14-SEP-1998;   98US-0100137P.
XX
PA   (DAND ) DANA FARBER CANCER INST INC.
XX
PI   Duke-Cohan JS,  Schlossman SF;
XX
DR   WPI; 2000-271373/23.
DR   P-PSDB; AAY70689.
XX
PT   Isolated nucleic acids encoding human attractin polypeptides useful for
PT   enhancing immune responses.
XX
PS   Claim 3; Fig 8; 120pp; English.
XX
CC   The patent discloses four forms of human attractin polypeptides which
CC   enhance immune response by promoting macrophage and monocyte spreading in
CC   the presence of T cells. These include soluble attractin-1 and -2 and
CC   membrane attractin-1 and -2. These various forms of attractin are encoded
CC   by alternatively spliced mRNA molecule transcribed from a single gene.
CC   The present sequence is a cDNA encoding soluble attractin-1 obtained from
CC   activated T cell and human foetal liver libraries. This sequence can be
CC   used to enhance immune response in immunosuppressed patients such as
CC   those undergoing chemo- and radio-therapy treatment for cancer or those
CC   suffering from common variable immunodeficiency syndrome. The proteins
CC   may also be used to screen modulators (agonists and antagonists) of
CC   immune responses which may also be used to regulate immune reactions.
CC   Attractin antibodies can be used to inhibit immune response in transplant
CC   recipients or patients afflicted with autoimmune disease
XX
SQ   Sequence 3597 BP; 951 A; 822 C; 898 G; 926 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                0              Length:          3597
Score:                    6612.00        Matches:         1191
Percent Similarity:       93.72%         Conservative:    2
Best Local Similarity:    93.56%         Mismatches:      5
Query Match:              83.54%         Indels:          76
DB:                       3              Gaps:            1

US-09-787-097-12 (1-1429) x AAZ52274 (1-3597)

Qy          1 MetValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAla 20
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 60

Qy         21 AlaLeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArg 40
                ||||||||||||||||||||||||||||||
Db         61 GCGCTCGCGGGCAGGAGCGGCGGGCCGCAC------------------------------ 90

Qy         41 ProGlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSerProProLeuArgProArg 60

Db         90 ------------------------------------------------------------ 90

Qy         61 LeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeuProCysGlu 80

Db         90 ------------------------------------------------------------ 90

Qy         81 AlaGluAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGlu 100

Db         90 ------------------------------------------------------------ 90

Qy        101 CysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCys 120
                              ||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 ------------TGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGC 138

Qy        121 ProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGly 140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 CCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGA 198

Qy        141 SerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrp 160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 TCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGG 258

Qy        161 LeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGlu 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 CTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAG 318

Qy        181 CysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAla 200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 TGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCT 378

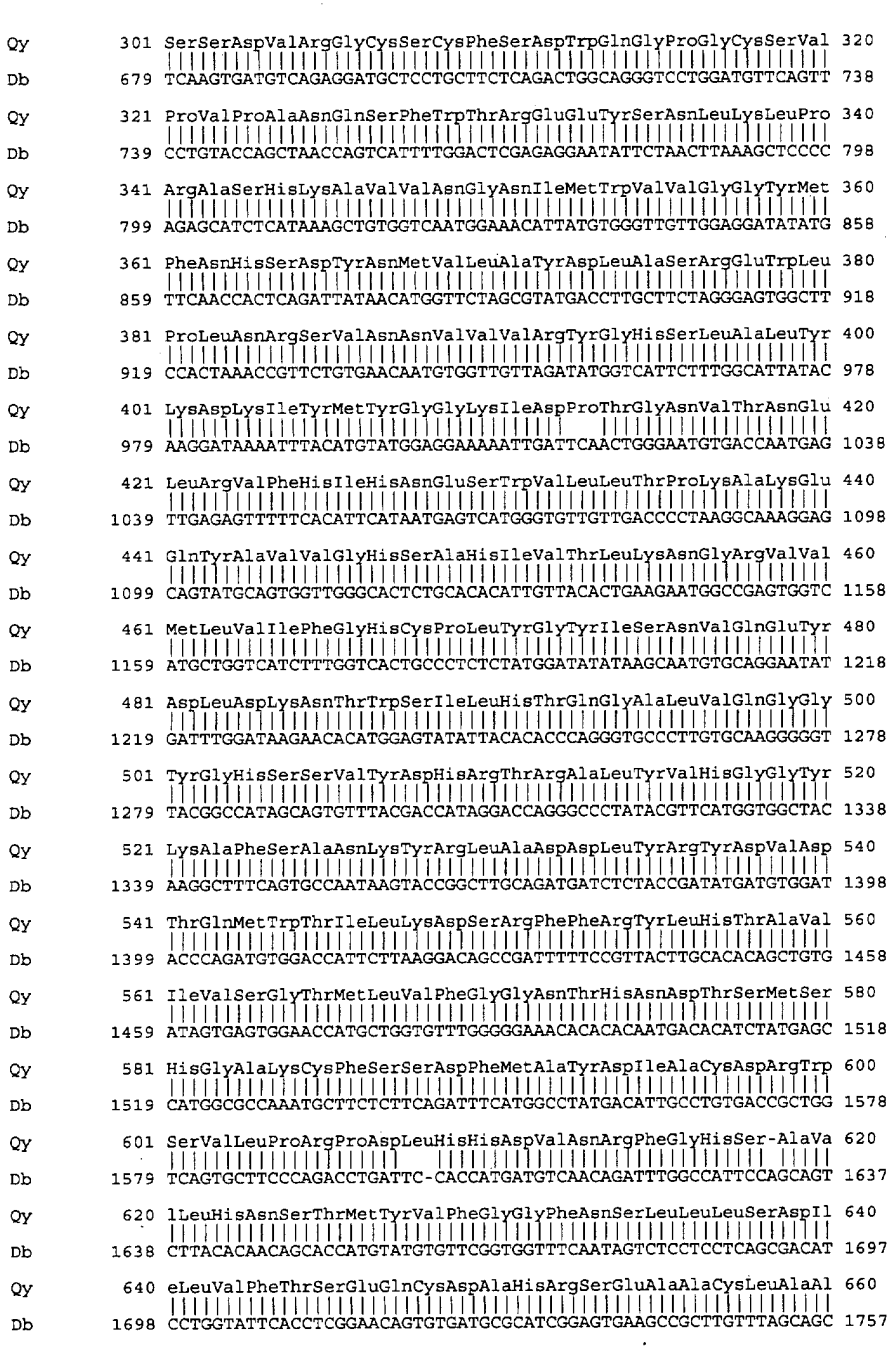
Qy        201 AlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGluValVal 220
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 GCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTT 438

Qy        221 AlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGly 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        439 GCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGA 498

Qy        241 PheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCys 260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 TTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGT 558

Qy        261 LysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsnTrpLysGlyGlu 280
                ||||||||||||||||||:::||||||||||||||||||||||||||||||||||||||||
Db        559 AAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAA 618

Qy        281 AlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsn 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        619 GCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAAT 678
```


Qy 301 SerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerVal 320
Db 679 TCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTT 738
Qy 321 ProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuPro 340
Db 739 CCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCC 798
Qy 341 ArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMet 360
Db 799 AGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATG 858
Qy 361 PheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeu 380
Db 859 TTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTT 918
Qy 381 ProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyr 400
Db 919 CCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATAC 978
Qy 401 LysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsnValThrAsnGlu 420
Db 979 AAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAATGAG 1038
Qy 421 LeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGlu 440
Db 1039 TTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGGAG 1098
Qy 441 GlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValVal 460
Db 1099 CAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTC 1158
Qy 461 MetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyr 480
Db 1159 ATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATAT 1218
Qy 481 AspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGly 500
Db 1219 GATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGT 1278
Qy 501 TyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyr 520
Db 1279 TACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTAC 1338
Qy 521 LysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAsp 540
Db 1339 AAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGAT 1398
Qy 541 ThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaVal 560
Db 1399 ACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTG 1458
Qy 561 IleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSer 580
Db 1459 ATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGC 1518
Qy 581 HisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrp 600
Db 1519 CATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGG 1578
Qy 601 SerValLeuProArgProAspLeuHisHisAspValAsnArgPheGlyHisSer-AlaVa 620
Db 1579 TCAGTGCTTCCCAGACCTGATTC-CACCATGATGTCAACAGATTTGGCCATTCCAGCAGT 1637
Qy 620 lLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIl 640
Db 1638 CTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACAT 1697
Qy 640 eLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAl 660
Db 1698 CCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGC 1757
Qy 660 aGlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLe 680
Db 1758 AGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCT 1817
Qy 680 uAlaThrAspGluGlnGluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAs 700
Db 1818 GGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGA 1877
Qy 700 pHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCy 720
Db 1878 CCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGACTG 1937
Qy 720 sHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSe 740
Db 1938 CCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTC 1997
Qy 740 rIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysTh 760
Db 1998 CATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAAGAC 2057
Qy 760 rSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCy 780
Db 2058 CAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTG 2117
Qy 780 sIleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLe 800
Db 2118 CATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTT 2177
Qy 800 uLysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAs 820
Db 2178 GAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAA 2237
Qy 820 nAlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuAr 840
Db 2238 TGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCG 2297
Qy 840 gIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuAr 860
Db 2298 AATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCG 2357
Qy 860 gLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLe 880
Db 2358 GAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACT 2417
Qy 880 uGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSe 900
Db 2418 ACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAG 2477
Qy 900 rThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluAr 920
Db 2478 TACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAG 2537
Qy 920 gProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGl 940
Db 2538 GCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGG 2597
Qy 940 yAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAs 960
Db 2598 AGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGA 2657
Qy 960 pSerAsnAlaTyrValAlaSerProProPheGlyGlnCysMetGluTrpTyrThrMetSe 980
Db 2658 CTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAG 2717
Qy 980 rThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPr 1000
Db 2718 CACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACC 2777
Qy 1000 oGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySe 1020
Db 2778 AGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTC 2837
Qy 1020 rTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnPr 1040

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2838 CTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCC 2897

Qy       1040 oLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCy 1060
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2898 CCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTG 2957

Qy       1060 sProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCy 1080
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2958 TCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTG 3017

Qy       1080 sGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspPr 1100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3018 TGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCC 3077

Qy       1100 oThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnTh 1120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3078 CACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACAC 3137

Qy       1120 rAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCy 1140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3138 CAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATG 3197

Qy       1140 sGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLe 1160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3198 TGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCT 3257

Qy       1160 uIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIl 1180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3258 TATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTAT 3317

Qy       1180 eAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSe 1200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3318 CAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTC 3377

Qy       1200 rLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGl 1220
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3378 CAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGG 3437

Qy       1220 yGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAs 1240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3438 AGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAA 3497

Qy       1240 nGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPh 1260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3498 TGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTT 3557

Qy       1260 eThrTrpProIleLysIleGlnIleAlaPheSerGln 1272
                ||||||||||||||||||||:::         |||
Db       3558 CACCTGGCCCATCAAAATTCAGGTGCAAACTGAACAA 3594

RESULT 10
AAS72657
ID   AAS72657 standard; cDNA; 3597 BP.
XX
AC   AAS72657;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #8461.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG08470.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 8461; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 3597 BP; 951 A; 822 C; 898 G; 926 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              0               Length:         3597
Score:                  6612.00         Matches:        1191
Percent Similarity:     93.72%          Conservative:   2
Best Local Similarity:  93.56%          Mismatches:     5
Query Match:            83.54%          Indels:         76
DB:                     5               Gaps:           1

US-09-787-097-12 (1-1429) x AAS72657 (1-3597)

Qy          1 MetValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 60

Qy         21 AlaLeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArg 40
              ||||||||||||||||||||||||||||||||
Db         61 GCGCTCGCGGGCAGGAGCGGCGGGCCGCAC---------------------------- 90

Qy         41 ProGlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSerProProLeuArgProArg 60

Db         90 ------------------------------------------------------------ 90

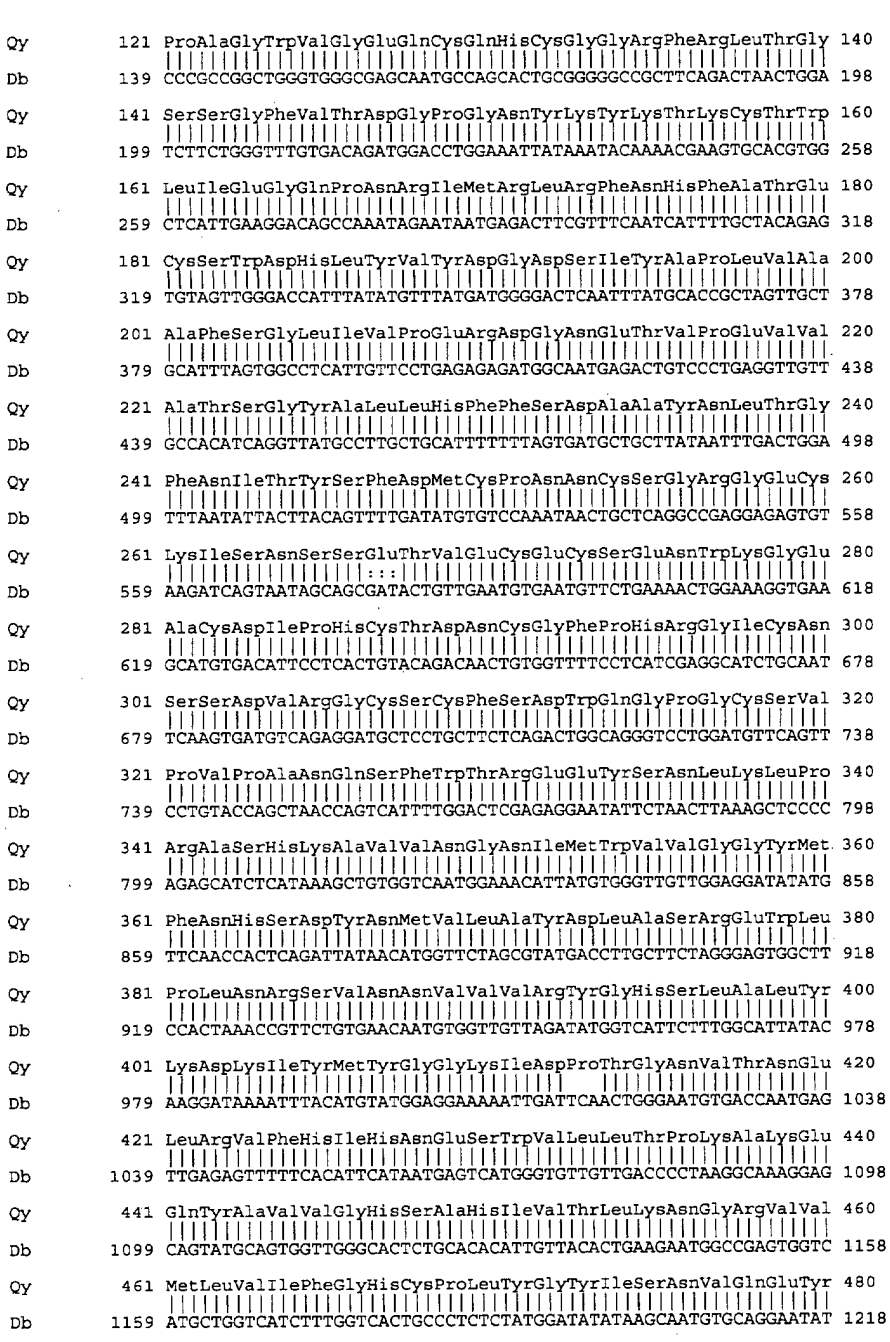
Qy         61 LeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeuProCysGlu 80

Db         90 ------------------------------------------------------------ 90

Qy         81 AlaGluAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGlu 100

Db         90 ------------------------------------------------------------ 90

Qy        101 CysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCys 120
                          ||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 -----------TGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGC 138
```


Qy 121 ProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGly 140
Db 139 CCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGA 198
Qy 141 SerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrp 160
Db 199 TCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGG 258
Qy 161 LeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGlu 180
Db 259 CTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAG 318
Qy 181 CysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAla 200
Db 319 TGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCT 378
Qy 201 AlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGluValVal 220
Db 379 GCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTT 438
Qy 221 AlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGly 240
Db 439 GCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGA 498
Qy 241 PheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCys 260
Db 499 TTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGT 558
Qy 261 LysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsnTrpLysGlyGlu 280
Db 559 AAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAA 618
Qy 281 AlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsn 300
Db 619 GCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAAT 678
Qy 301 SerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerVal 320
Db 679 TCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTT 738
Qy 321 ProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuPro 340
Db 739 CCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCC 798
Qy 341 ArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMet 360
Db 799 AGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATG 858
Qy 361 PheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeu 380
Db 859 TTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTT 918
Qy 381 ProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyr 400
Db 919 CCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATAC 978
Qy 401 LysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsnValThrAsnGlu 420
Db 979 AAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAATGAG 1038
Qy 421 LeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGlu 440
Db 1039 TTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGGAG 1098
Qy 441 GlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValVal 460
Db 1099 CAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTC 1158
Qy 461 MetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyr 480
Db 1159 ATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATAT 1218
Qy 481 AspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGly 500
Db 1219 GATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGT 1278
Qy 501 TyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyr 520
Db 1279 TACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTAC 1338
Qy 521 LysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAsp 540
Db 1339 AAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGAT 1398
Qy 541 ThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaVal 560
Db 1399 ACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTG 1458
Qy 561 IleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSer 580
Db 1459 ATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGC 1518
Qy 581 HisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrp 600
Db 1519 CATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGG 1578
Qy 601 SerValLeuProArgProAspLeuHisHisAspValAsnArgPheGlyHisSer-AlaVa 620
Db 1579 TCAGTGCTTCCCAGACCTGATTC-CACCATGATGTCAACAGATTTGGCCATTCCAGCAGT 1637
Qy 620 lLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIl 640
Db 1638 CTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACAT 1697
Qy 640 eLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAl 660
Db 1698 CCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGC 1757
Qy 660 aGlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLe 680
Db 1758 AGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCT 1817
Qy 680 uAlaThrAspGluGlnGluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAs 700
Db 1818 GGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGA 1877
Qy 700 pHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCy 720
Db 1878 CCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGACTG 1937
Qy 720 sHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSe 740
Db 1938 CCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTC 1997
Qy 740 rIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysTh 760
Db 1998 CATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAAGAC 2057
Qy 760 rSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCy 780
Db 2058 CAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTG 2117
Qy 780 sIleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLe 800
Db 2118 CATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTT 2177
Qy 800 uLysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAs 820
Db 2178 GAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAA 2237
Qy 820 nAlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuAr 840
Db 2238 TGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCG 2297
Qy 840 gIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuAr 860

```
Db       2298 AATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCG 2357

Qy        860 gLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLe 880

Db       2358 GAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACT 2417

Qy        880 uGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSe 900

Db       2418 ACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAG 2477

Qy        900 rThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluAr 920

Db       2478 TACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAG 2537

Qy        920 gProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGl 940

Db       2538 GCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGG 2597

Qy        940 yAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAs 960

Db       2598 AGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGA 2657

Qy        960 pSerAsnAlaTyrValAlaSerProProheGlyGlnCysMetGluTrpTyrThrMetSe 980

Db       2658 CTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAG 2717

Qy        980 rThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPr 1000

Db       2718 CACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACC 2777

Qy       1000 oGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySe 1020

Db       2778 AGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTC 2837

Qy       1020 rTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnPr 1040

Db       2838 CTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCC 2897

Qy       1040 oLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCy 1060

Db       2898 CCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTG 2957

Qy       1060 sProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCy 1080

Db       2958 TCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTG 3017

Qy       1080 sGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspPr 1100

Db       3018 TGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCC 3077

Qy       1100 oThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnTh 1120

Db       3078 CACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACAC 3137

Qy       1120 rAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCy 1140

Db       3138 CAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATG 3197

Qy       1140 sGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLe 1160

Db       3198 TGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCT 3257

Qy       1160 uIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIl 1180

Db       3258 TATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTAT 3317

Qy       1180 eAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSe 1200

Db       3318 CAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTC 3377

Qy       1200 rLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGl 1220

Db       3378 CAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGG 3437

Qy       1220 yGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAs 1240

Db       3438 AGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAA 3497

Qy       1240 nGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPh 1260

Db       3498 TGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTT 3557

Qy       1260 eThrTrpProIleLysIleGlnIleAlaPheSerGln 1272

Db       3558 CACCTGGCCCATCAAAATTCAGGTGCAAACTGAACAA 3594
```

RESULT 11
ADD70997
ID ADD70997 standard; DNA; 3597 BP.
XX
AC ADD70997;
XX
DT 15-JAN-2004 (first entry)
XX
DE Human attractin gene SEQ ID NO:1.
XX
KW liver cancer; chronic hepatitis; cirrhosis; liver disease; hepatotropic;
KW cytostatic; gene therapy; human; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2003061564-A2.
XX
PD 31-JUL-2003.
XX
PF 20-DEC-2002; 2002WO-US040718.
XX
PR 21-DEC-2001; 2001US-0341815P.
PR 31-DEC-2001; 2001US-0343185P.
XX
PA (GENE-) GENE LOGIC INC.
PA (LGBI-) LG BIOMEDICAL INST.
XX
PI Koh SS, Liu Q, Chung H, Zeng W, Lee B, Yeramilli S, Song SY;
XX
DR WPI; 2003-663343/62.
XX
PT Diagnosing liver cancer cells, useful for treating liver cancer
PT associated with chronic hepatitis or cirrhosis comprises detecting the
PT level of expression in a tissue sample of one or more genes associated
PT with cancerous liver tissues.
XX
PS Claim 1; SEQ ID NO 1; 176pp; English.
XX
CC The present invention describes a method for diagnosing liver cancer
CC cells comprising detecting the level of expression in a tissue sample of
CC one or more genes given in the specification (see ADD70997 to ADD71105),
CC where differential expression of the genes is indicative of liver cancer.
CC Also described: (1) detecting the progression of liver cancer in a
CC patient; (2) monitoring the treatment of a patient with liver cancer; (3)
CC treating a patient with liver cancer; (4) typing a liver disease in a
CC patient; (5) detecting the presence or progression of liver cancer in a
CC patient with chronic hepatitis or cirrhosis; (6) differentiating liver
CC cancer related to chronic hepatitis from liver cancer related to
CC cirrhosis; (7) screening for an agent capable of modulating the onset or
CC progression of liver cancer; (8) a composition comprising at least two
CC oligonucleotides comprising a sequence that specifically hybridises to
CC any of the genes; (9) a solid support comprising the at least two
CC oligonucleotides; (10) a computer system comprising a database containing
CC information identifying the level in liver tissue of a set of genes; (11)
CC a method for using the computer system to present information identifying
CC the expression level in tissue or cell of any of the genes; and (12) a
CC therapeutic agent for slowing or halting the progression of liver cancer.
CC The methods are useful for treating liver cancer associated with chronic
CC hepatitis or cirrhosis. The present sequence represents a specifically

CC claimed human gene sequence which is used in the exemplification of the
CC present invention.
XX
SQ Sequence 3597 BP; 951 A; 822 C; 898 G; 926 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	3597
Score:	6612.00	Matches:	1191
Percent Similarity:	93.72%	Conservative:	2
Best Local Similarity:	93.56%	Mismatches:	5
Query Match:	83.54%	Indels:	76
DB:	9	Gaps:	1

US-09-787-097-12 (1-1429) x ADD70997 (1-3597)

```
Qy          1 MetValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAla 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 60

Qy         21 AlaLeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArg 40
              ||||||||||||||||||||||||||||||
Db         61 GCGCTCGCGGGCAGGAGCGGCGGGCCGCAC---------------------------- 90

Qy         41 ProGlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSerProProLeuArgProArg 60

Db         90 ------------------------------------------------------------ 90

Qy         61 LeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeuProCysGlu 80

Db         90 ------------------------------------------------------------ 90

Qy         81 AlaGluAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGlu 100

Db         90 ------------------------------------------------------------ 90

Qy        101 CysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCys 120
                           ||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 ------------TGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGC 138

Qy        121 ProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGly 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 CCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGA 198

Qy        141 SerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrp 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 TCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGG 258

Qy        161 LeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 CTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAG 318

Qy        181 CysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAla 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 TGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCT 378

Qy        201 AlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGluValVal 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 GCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTT 438

Qy        221 AlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGly 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        439 GCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGA 498

Qy        241 PheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCys 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 TTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGT 558

Qy        261 LysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsnTrpLysGlyGlu 280
              ||||||||||||||||||:::||||||||||||||||||||||||||||||||||||||||
Db        559 AAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAA 618

Qy        281 AlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsn 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        619 GCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAAT 678

Qy        301 SerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerVal 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        679 TCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTT 738

Qy        321 ProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuPro 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        739 CCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCC 798

Qy        341 ArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMet 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        799 AGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATG 858

Qy        361 PheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeu 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        859 TTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTT 918

Qy        381 ProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        919 CCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATAC 978

Qy        401 LysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsnValThrAsnGlu 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        979 AAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAATGTGACCAATGAG 1038

Qy        421 LeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGlu 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1039 TTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGGAG 1098

Qy        441 GlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValVal 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1099 CAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTC 1158

Qy        461 MetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyr 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1159 ATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATAT 1218

Qy        481 AspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGly 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1219 GATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGT 1278

Qy        501 TyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyr 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1279 TACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTAC 1338

Qy        521 LysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAsp 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1339 AAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGAT 1398

Qy        541 ThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaVal 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1399 ACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTG 1458

Qy        561 IleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSer 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1459 ATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGC 1518

Qy        581 HisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrp 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1519 CATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGG 1578

Qy        601 SerValLeuProArgProAspLeuHisHisAspValAsnArgPheGlyHisSer-AlaVa 620
              ||||||||||||||||||||  ||||||||||||||||||||||||||||| |||||
Db       1579 TCAGTGCTTCCCAGACCTGATTC-CACCATGATGTCAACAGATTTGGCCATTCCAGCAGT 1637

Qy        620 lLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIl 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1638 CTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACAT 1697

Qy        640 eLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAl 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1698 CCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGC 1757
```

```
Qy       660 aGlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLe 680
Db      1758 AGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCT 1817

Qy       680 uAlaThrAspGluGlnGluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAs 700
Db      1818 GGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGA 1877

Qy       700 pHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCy 720
Db      1878 CCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGACTG 1937

Qy       720 sHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSe 740
Db      1938 CCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTC 1997

Qy       740 rIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysTh 760
Db      1998 CATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAAGAC 2057

Qy       760 rSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCy 780
Db      2058 CAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTG 2117

Qy       780 sIleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLe 800
Db      2118 CATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTT 2177

Qy       800 uLysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAs 820
Db      2178 GAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAA 2237

Qy       820 nAlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuAr 840
Db      2238 TGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCG 2297

Qy       840 gIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuAr 860
Db      2298 AATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCG 2357

Qy       860 gLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLe 880
Db      2358 GAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACT 2417

Qy       880 uGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSe 900
Db      2418 ACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAG 2477

Qy       900 rThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluAr 920
Db      2478 TACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAG 2537

Qy       920 gProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGl 940
Db      2538 GCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGG 2597

Qy       940 yAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAs 960
Db      2598 AGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGA 2657

Qy       960 pSerAsnAlaTyrValAlaSerProProPheGlyGlnCysMetGluTrpTyrThrMetSe 980
Db      2658 CTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAG 2717

Qy       980 rThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPr 1000
Db      2718 CACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACC 2777

Qy      1000 oGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySe 1020
Db      2778 AGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTC 2837

Qy      1020 rTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnPr 1040
Db      2838 CTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCC 2897

Qy      1040 oLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCy 1060
Db      2898 CCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTG 2957

Qy      1060 sProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCy 1080
Db      2958 TCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTG 3017

Qy      1080 sGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspPr 1100
Db      3018 TGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCC 3077

Qy      1100 oThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnTh 1120
Db      3078 CACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACAC 3137

Qy      1120 rAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCy 1140
Db      3138 CAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATG 3197

Qy      1140 sGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLe 1160
Db      3198 TGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCT 3257

Qy      1160 uIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIl 1180
Db      3258 TATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTAT 3317

Qy      1180 eAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSe 1200
Db      3318 CAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTC 3377

Qy      1200 rLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGl 1220
Db      3378 CAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGG 3437

Qy      1220 yGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAs 1240
Db      3438 AGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAA 3497

Qy      1240 nGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPh 1260
Db      3498 TGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTT 3557

Qy      1260 eThrTrpProIleLysIleGlnIleAlaPheSerGln 1272
Db      3558 CACCTGGCCCATCAAAATTCAGGTGCAAACTGAACAA 3594
```

RESULT 12
AAS72660
ID AAS72660 standard; cDNA; 3490 BP.
XX
AC AAS72660;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #8464.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX

PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG08473.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 8464; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 3490 BP; 931 A; 772 C; 858 G; 929 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	3490
Score:	5622.00	Matches:	1070
Percent Similarity:	92.77%	Conservative:	8
Best Local Similarity:	92.08%	Mismatches:	18
Query Match:	71.03%	Indels:	67
DB:	5	Gaps:	6

US-09-787-097-12 (1-1429) x AAS72660 (1-3490)

```
Qy        137 ArgLeuThrGlySerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThr 156
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 AGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACG 81

Qy        157 LysCysThrTrpLeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHis 176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 AAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCAT 141

Qy        177 PheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAla 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 TTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCA 201

Qy        197 ProLeuValAlaAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrVal 216
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 CCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTC 261

Qy        217 ProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyr 236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 CCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTAT 321

Qy        237 AsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGly 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 AATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGC 381

Qy        257 ArgGlyGluCysLysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsn 276
              ||||||||||||||||||||||||||||:::||||||||||||||||||||||||||||||
Db        382 CGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAAC 441

Qy        277 TrpLysGlyGluAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArg 296
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        442 TGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGA 501

Qy        297 GlyIleCysAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyPro 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        502 GGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCT 561

Qy        317 GlyCysSerValProValProAlaAsnGlnSerPheTrp-ThrArgGluGluTyrSerAs 336
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        562 GGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGGACTCGAGAGGAATATTCTAA 621

Qy        336 n-LeuLysLeu-ProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpVal 355
              | |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        622 CTTTAAAGCTCCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTT 681

Qy        356 ValGlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAla 375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        682 GTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCT 741

Qy        376 SerArgGluTrpLeuProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHis 395
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        742 TCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCAT 801

Qy        396 SerLeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGly 415
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        802 TCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGG 861

Qy        416 AsnValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThr 435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        862 AATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACC 921

Qy        436 ProLysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLys 455
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        922 CCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAG 981

Qy        456 AsnGlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSer 475
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        982 AATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGC 1041

Qy        476 AsnValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAla 495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1042 AATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCC 1101

Qy        496 LeuValGlnGlyGlyTyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyr 515
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1102 CTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATAC 1161

Qy        516 ValHisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyr 535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1162 GTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTAC 1221

Qy        536 ArgTyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyr 555
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1222 CGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTAC 1281

Qy        556 LeuHisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsn 575
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1282 TTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGAGGAAACACACACAAT 1341

Qy        576 AspThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIle 595
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1342 GACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATT 1401

Qy        596 AlaCysAspArgTrpSerValLeuProArgProAspLeuHisHisAspValAsnArgPhe 615
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1402 GCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTT 1461
```

```
Qy       616 GlyHisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeu 635
Db      1462 GGCCATTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTC 1521

Qy       636 LeuLeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAla 655
Db      1522 CTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCC 1581

Qy       656 AlaCysLeuAlaAlaGlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCys 675
Db      1582 GCTTGTTTAGCAGCAGGACCTGGTATTCGGTGTCT---------------------------- 1616

Qy       676 IleSerTrpAlaLeuAlaThrAspGluGlnGluGluLysLeuLysSerGluCysPheSer 695
Db      1616 ------------------------------------------------------------ 1616

Qy       696 LysArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAla 715
Db      1617 ---------CTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCC 1667

Qy       716 AsnThrAsnAspCysHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSer 735
Db      1668 AACACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCA 1727

Qy       736 GluGlyGlnIleSerIlePheArg-TyrGluAsnCysProLysAspAsnProMetTyrTy 755
Db      1728 GAAGGC--------------CAGGTATGAGAATTGCCCCAAGGATAACCCCATGTACTA 1772

Qy       755 rCysAsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluPr 775
Db      1773 CTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCC 1832

Qy       775 oArgAsnGlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuVa 795
Db      1833 CCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGT 1892

Qy       795 lGlyAsnSerCysLeuLysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPh 815
Db      1893 TGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTT 1952

Qy       815 eCysArgAsnHisAsnAlaLeuLeuAlaSerLeuThrThrGlnLysLysVal-GluPheV 835
Db      1953 CTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGGAATTTG 2012

Qy       835 alLeuLys--GlnLeuArgIleMetGln---SerSerGlnSerMetSerLysLeuThrLe 853
Db      2013 TCCTTAAGGCAGGCTGCGAATAATGCAGGTCATCTCAGAGGCATGTCCAAGGCTCACCTT 2072

Qy       853 uThrProTrp---ValGlyLeuArgLysIleAsn-ValSerTyrTrpCysTrpGluAspM 872
Db      2073 AACCCCATGGGGTCGGGCCTTCGGGAAGATCAATTGTGTCCTACTGGTGCTGGGAAGATA 2132

Qy       872 etSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyP 892
Db      2133 TGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGAT 2192

Qy       892 heCysGly---IleLeuSerGluProSer--ThrArgGlyLeuLysAlaAlaThrCysIl 910
Db      2193 TCTGTGGGAATTTTATTCAGGAACCCAGTTACTTCGGGGACTGAAGGCTGCAACCTGCAT 2252

Qy       910 e-AsnProLeu-AsnGlySerValCysGlu-ArgPro-AlaAsnHisSer-AlaLysGln 928
Db      2253 TCAACCCACTTCAATGGTAGTGTCTGTGAAAAGGCCTGGCAAACCACAGTGGTTAAGCAG 2312

Qy       929 CysArgThrPro---CysAlaLeuArg-ThrAlaCysGlyAspCysThrSerGlySerSe 947
Db      2313 TGCCGGACACCATGGTGCCCTTTGAGGGACAGCATGTGGAGATTGCACCAGCGGCAGCTC 2372

Qy       947 rGluCys-MetTrpCysSerAsnMetLys-GlnCysValAspSerAsnAlaTyrValAla 966
Db      2373 TGAGTGGCATGTGGTGCAGCAACATGAAGGCAGTGTGTGGACTCCAATGCCTATGTGGCC 2432

Qy       967 SerPheProPhe-GlyGlnCysMetGluTrpTyrThrMet-SerThrCysProProGluA 986
Db      2433 TCCTTCCCTTTTGGGCCAGTGTATGGAATGGTATACGATGAGGCACCTGCCCCCCTGAAA 2492

Qy       986 snCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPro-GlyCysGlyTrpCys 1005
Db      2493 ATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGGCTGTGGCTGGTGT 2552

Qy      1006 ThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyProVal 1025
Db      2553 ACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTG 2612

Qy      1026 LysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerSer 1045
Db      2613 AAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGC 2672

Qy      1046 MetCysLeuGluAspSerArgTyrAsnTrp-SerPheIle-HisCysProAlaCysGlnC 1065
Db      2673 ATGTGTCTAGAGGACAGCAGATACAACTGGGTCTTTCATTTCACTGTCCAGTTTGCCAAT 2732

Qy      1065 ysAsn-GlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThr 1084
Db      2733 GGAACGGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACC 2792

Qy      1085 ThrGlyLysHisCysGluThrCysIleSerGlyPheTyr-GlyAspProThrAsn-GlyG 1104
Db      2793 ACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGGTGATCCCACCAATTGGAG 2852

Qy      1104 lyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyL 1124
Db      2853 GGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCA 2912

Qy      1124 ysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGluA 1144
Db      2913 AGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAA 2972

Qy      1144 snArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrG 1164
Db      2973 ATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATC 3032

Qy      1164 lnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIleAsnPheValA 1184
Db      3033 AGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGG 3092

Qy      1184 laThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheA 1204
Db      3093 CTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCA 3152

Qy      1204 snLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMetP 1224
Db      3153 ACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGC 3212

Qy      1224 roValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPheA 1244
Db      3213 CTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTG 3272

Qy      1244 spPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpProI 1264
Db      3273 ATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCA 3332

Qy      1264 leLysIleGlnIle 1268
Db      3333 TCAAAATTCAGGTG 3346
```

RESULT 13
ABZ58709
ID ABZ58709 standard; cDNA; 4140 BP.
XX
AC ABZ58709;
XX
DT 14-APR-2003 (first entry)
XX
DE Human homologue of attractin/mahogany (HAM) polypeptide encoding cDNA.

```
XX
KW   HAM; homologue of attractin/mahogany; immunosuppressive; cytostatic;
KW   antiinflammatory; cardiant; osteopathic; gene therapy; human; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .4140
FT                   /*tag= a
FT                   /product= "HAM"
FT                   /note= "homologue of attractin/mahogany"
FT   sig_peptide     1. .180
FT                   /*tag= b
FT   mat_peptide     181. .4137
FT                   /*tag= c
XX
PN   WO200297120-A1.
XX
PD   05-DEC-2002.
XX
PF   23-MAY-2002; 2002WO-US016391.
XX
PR   25-MAY-2001; 2001US-0293608P.
PR   24-SEP-2001; 2001US-0324626P.
XX
PA   (IMMV ) IMMUNEX CORP.
XX
PI   Anderson DM;
XX
DR   WPI; 2003-140486/13.
DR   P-PSDB; ABP71087.
XX
PT   New Homologue of Attractin/Mahagony (HAM) polypeptide, useful for
PT   treating HAM-associated disorder consisting of inflammatory, autoimmune,
PT   cell proliferative or cardiovascular disorders.
XX
PS   Claim 14; Fig 1A-D; 89pp; English.
XX
CC   The invention relates to Homologue of Attractin/Mahogany (HAM)
CC   polypeptides and encoding polynucleotides. The HAM polypeptides can be
CC   expressed by standard recombinant methodology. The HAM polypeptides are
CC   useful for treating HAM-associated disorder consisting of inflammatory,
CC   autoimmune, graft-versus-host, neurological, myelination, cell
CC   proliferative, cardiovascular, haematologic, liver, metabolic, weight or
CC   bone disorder. The present sequence represents a cDNA encoding a human
CC   HAM polypeptide
XX
SQ   Sequence 4140 BP; 1209 A; 816 C; 941 G; 1174 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.99e-263       Length:         4140
Score:                  4637.50         Matches:        843
Percent Similarity:     73.88%          Conservative:   209
Best Local Similarity:  59.20%          Mismatches:     300
Query Match:            58.59%          Indels:         72
DB:                     7               Gaps:           23

US-09-787-097-12 (1-1429) x ABZ58709 (1-4140)

Qy         28 GlyProHisTrpAspTrpAspValThrArgAlaGlyArgProGly---------------- 42
              ||||||||||||           |||      ||||||
Db         18 GGCCCGCACTGG---------TACCCCGCAGCCAGCGGCCCCGGGGGTGTGGAGGGCTCG 68

Qy         43 LeuGlyAlaGlyLeuArgLeuProArgLeuLeuSerProProLeuArgProArgLeuLeu 62
                 |||          |||      |||||||||           |||   :::|||
Db         69 GCCGGCGGGCGGCGGCGGCGGGGGCGCCTCCTCCTGGCTGCTGGACGGGAACAGCTGGCT 128

Qy         63 LeuLeuLeuLeuLeuLeuProProProLeu-LeuLeuLeuLeuLeuProCysGluAlaGl 82
                 :::|||   |||          |||||| |||   |||
Db        129 GCTGTGCTATGGCTT---------CCTCTACCTGGCGCTCTAC----------------- 162

Qy         82 uAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGluCysAs 102
                                          :::      :::::::::|||   |||::
```

```
Db        163 ------------------------------GCGCAGGTGTCCCAGTCCAAGCCGTGCGA 191

Qy        102 pArg------ProCysValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCy 120
                :|||          |||   :::   |||||||||           |||      |||:::||
Db        192 GAGGACCGGCTCCTGCTTCTCG---GGCCGCTGTGTCAACTCCACC------TGCCTCTG 242

Qy        120 sProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGl 140
                |     ||||||||||||:::||||||||||||||||    |||||||||:::||||||
Db        243 CGACCCGGGCTGGGTGGGGGACCAGTGCCAGCACTGCCAGGGCAGGTTCAAGTTAACAGA 302

Qy        140 ySerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTr 160
                  ||||||::::::|||||||||||||   |||||||||||||||||||||||||||||
Db        303 ACCTTCTGGATATTTAACAGATGGCCCAATTAACTATAAATATAAAACTAAATGTACTTG 362

Qy        160 pLeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGl 180
                |||||||||||||    ||||||   ::::::|||||||||||||||||||||||||||||
Db        363 GCTCATTGAAGGCTATCCAAATGCAGTGTTAAGATTAAGATTCAATCATTTTGCTACAGA 422

Qy        180 uCysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAl 200
                |||||||||||||||:::|||||||||||||||||||||||||||||||||||:::||
Db        423 ATGTAGCTGGGATCATATGTATGTTTATGATGGAGATTCAATATATGCACCTTTAATAGC 482

Qy        200 aAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGluValVa 220
                |     ||||||||||||||||||||||     ||||||||||||||||||||||||||||
Db        483 TGTACTTAGTGGTTTGATAGTCCCTGAAATAAGGGGCAATGAAACTGTGCCTGAAGTTGT 542

Qy        220 lAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGl 240
                | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 TACTACATCTGGCTATGCACTGTTACATTTTTTTAGTGATGCTGCGTATAATCTAACTGG 602

Qy        240 yPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCy 260
                ||||||||||    ||||||   :::   ||||||||||||||||||||   |||:::||
Db        603 TTTCAACATTTTCTATTCAATCAATTCTTGTCCTAACAATTGCTCTGGTCATGGGAAGTG 662

Qy        260 sLysIleSerAsnSer---SerGluThrValGluCysGluCysSerGluAsnTrpLysGl 279
                |     |||   |||          |||   |||||||||    :::   ||||||||
Db        663 TACAACTAGTGTCTCTGTTCCAAGTCAAGTATATTGTGAATGTGATAAATACTGGAAGGG 722

Qy        279 yGluAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCy 299
                |||||||||||||||||:::|||      |||||||||   |||      |||   ||
Db        723 TGAAGCTTGTGATATTCCTTACTGTAAAGCCAATTGCGGCAGTCCAGATCACGGTTACTG 782

Qy        299 sAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSe 319
                |:::   :::      :::   |||   |||          ||||||||||    |||||
Db        783 TGACCTGACTGGAGAAAAATTATGTGTCTGCAATGATAGTTGGCAAGGTCCTGATTGTTC 842

Qy        319 rValProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeuLys-- 338
                |:::   ||||||:::   :::|||:::|||          |||:::|||
Db        843 TTTGAATGTTCCCTCTACTGAGTCTTACTGGATT---------CTGCCAAACGTTAAACC 893

Qy        339 ------------LeuProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTr 354
                         :::   |||||||||||||||||||:::::|||      |||||
Db        894 CTTCAGTCCTTCTGTAGGTCGGGCTTCACATAAAGCAGTTTTACACGGGAAATTTATGTG 953

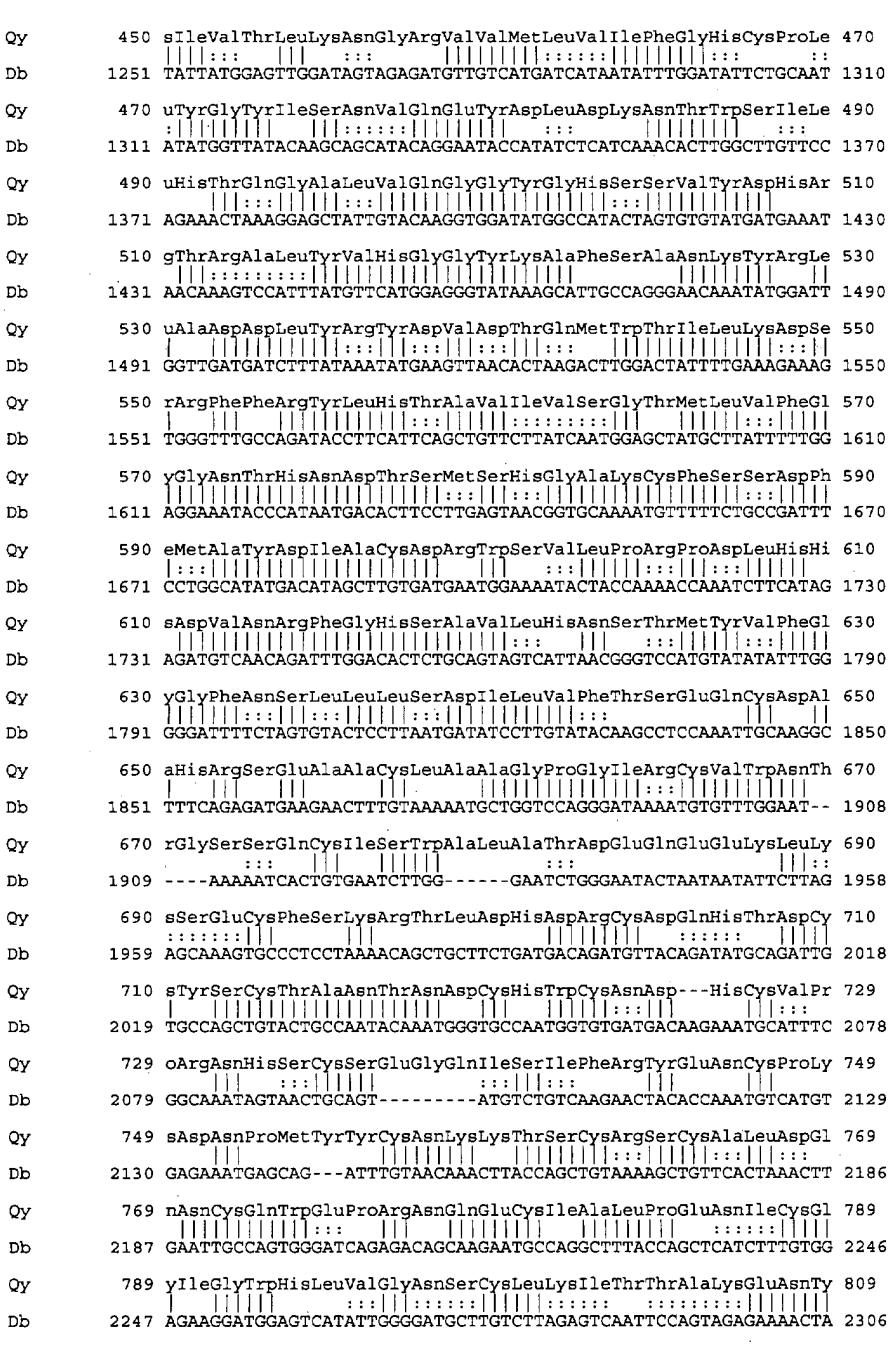
Qy        354 pValValGlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLe 374
                ||||:::|||||||||   ||||||:::|||   :::   |||||||||   |||:::||
Db        954 GGTGATTGGTGGATATACTTTTAACTACAGTTCTTTTCAAATGGTCCTAAATTACAATTT 1013

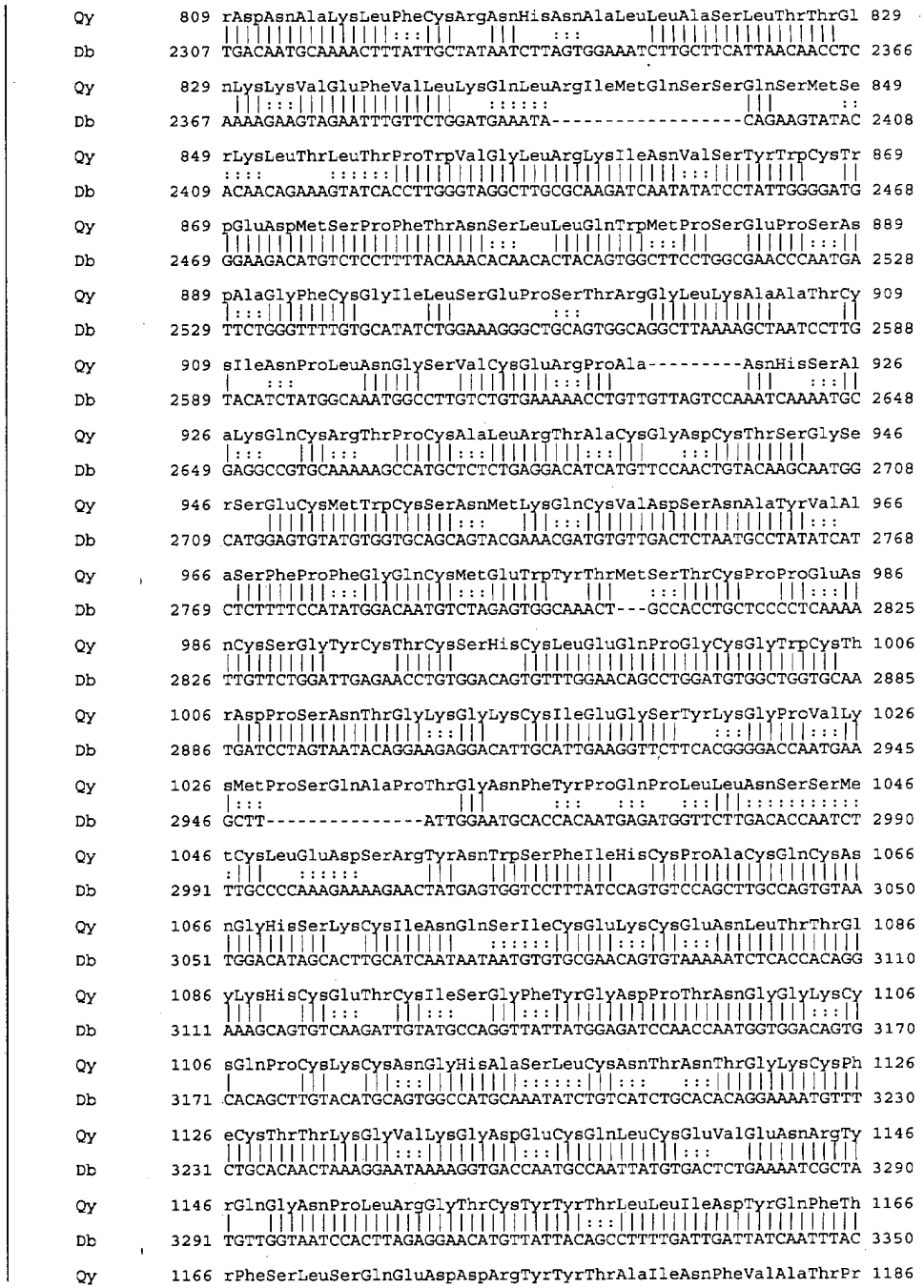
Qy        374 uAlaSerArgGluTrpLeuProLeuAsnArgSerValAsnAsnValValValArgTyrGl 394
                |   |||      |||      :::                        :::   ||||||||
Db       1014 AGAAAGCAGTATATGG---AATGTAGGAACTCCATCAAGGGGACCTCTCCAGAGATATGG 1070

Qy        394 yHisSerLeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLysIleAspProTh 414
                |||||||||||||||||:::::   |||:::|||||||||||:::|||:::
Db       1071 ACACTCTCTTGCTTTATATCAGGAAAACATCTTTATGTATGGAGGCAGAATTGAAACAAA 1130

Qy        414 r---GlyAsnValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLe 433
                    ||||||||||||:::||||||   |||||:::||||||:::::|||||
Db       1131 TGATGGCAATGTCACAGATGAATTATGGGTTTTTAACATACATAGTCAGTCATGGAGTAC 1190

Qy        433 uLeuThrProLysAla---------LysGluGlnTyrAlaValValGlyHisSerAlaHi 450
                |||||                  :::|||||||||||   ||||||||||||||
Db       1191 AAAAACTCCTACTGTTCTTGGACATGGTCAGCAGTATGCTGTGGAGGGACATTCAGCACA 1250
```





```
Db       3351 CTTCAGCTTATTACAGGAAGATGATCGCCACCATACTGCCATAAACTTTATAGCAAACCC 3410

Qy       1186 oAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeuAs 1206

Db       3411 AGAACAGTCGAACAAAAATCTGGATATATCAATTAATGCATCAAACAACTTTAATCTCAA 3470

Qy       1206 nIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProValVa 1226

Db       3471 CATTACGTGGTCTGTCGGTTCAACAGCTGGAACAATATCTGGGGAAGAGACTTCTATAGT 3530

Qy       1226 lSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPheAspPheAr 1246

Db       3531 TTCCAAGAATAATATAAAGGAATACAGAGATAGTTTTTCCTATGAAAAATTTAACTTTAG 3590

Qy       1246 gAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpProIleLysIl 1266

Db       3591 AAGCAATCCTAACATTACGTTCTATGTGTACGTCAGCAACTTTTCCTGGCCTATTAAAAT 3650

Qy       1266 eGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePheValThrPh 1286

Db       3651 ACAGATTGCATTCTCACAACACAATACAATCATGGACCTTGTGCAGTTTTTTGTCACCTT 3710

Qy       1286 ePheSerCysPheLeuSerLeuLeuLeuValAlaAlaValValTrpLysIleLysGlnSe 1306

Db       3711 CTTCAGTTGTTTCCTATCCTTATTGCTGGTGGCTGCTGTGGTATGGAAGATCAAACAAAC 3770

Qy       1306 rCysTrpAlaSerArgArgArgGluGlnLeuLeuArgGluMetGlnGlnMetAlaSerAr 1326

Db       3771 TTGTTGGGCTTCTCGACGGAGAGAGCAACTGCTTCGAGAACGACAGCAGATGGCCAGCCG 3830

Qy       1326 gProPheAlaSerValAsnValAlaLeuGluThrAspGluGluProProAspLeuIleGl 1346

Db       3831 TCCCTTTGCTTCTGTTGATGTAGCTCTGGAAGTGGGAGCTGAACAAACAGAGTTTCTGCG 3890

Qy       1346 yGlySerIleLysThrValProLysProIleAlaLeuGluProCysPheGlyAsnLysAl 1366

Db       3891 AGGGCCATTAGAGGGGGCACCCAAGCCAATTGCCATTGAACCATGTGCTGGGAACAGAGC 3950

Qy       1366 aAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyGlyIleProProProGl 1386

Db       3951 TGCTGTTCTGACTGTGTTTCTTTGTCTACCACGAGGATCATCAGGTGCCCCTCCCCCTGG 4010

Qy       1386 yGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetProIleVa 1406

Db       4011 GCAGTCAGGCCTTGCAATTGCCAGTGCCCTAATAGATATTTCACAACAGAAAGCTTCAGA 4070

Qy       1406 lTyrLysGluLysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnProGl 1426

Db       4071 TAGTAAAGATAAGACTTCTGGAGTCCGGAATCGAAAACACCTTTCAACACGTCAA---GG 4127

Qy       1426 yThrCysIle 1429

Db       4128 AACTTGTGTC 4137
```

RESULT 14
ABZ58710
ID ABZ58710 standard; cDNA; 4134 BP.
XX
AC ABZ58710;
XX
DT 14-APR-2003 (first entry)
XX
DE Murine homologue of attractin/mahogany (HAM) polypeptide encoding cDNA.
XX
KW HAM; homologue of attractin/mahogany; immunosuppressive; cytostatic;
KW antiinflammatory; cardiant; osteopathic; gene therapy; mouse; gene; ss.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 1. .4134
FT /*tag= a
FT /product= "HAM"
FT /note= "homologue of attractin/mahogany"
XX
PN WO200297120-A1.
XX
PD 05-DEC-2002.
XX
PF 23-MAY-2002; 2002WO-US016391.
XX
PR 25-MAY-2001; 2001US-0293608P.
PR 24-SEP-2001; 2001US-0324626P.
XX
PA (IMMV) IMMUNEX CORP.
XX
PI Anderson DM;
XX
DR WPI; 2003-140486/13.
DR P-PSDB; ABP71102.
XX
PT New Homologue of Attractin/Mahagony (HAM) polypeptide, useful for
PT treating HAM-associated disorder consisting of inflammatory, autoimmune,
PT cell proliferative or cardiovascular disorders.
XX
PS Claim 14; Fig 5A-D; 89pp; English.
XX
CC The invention relates to Homologue of Attractin/Mahagony (HAM)
CC polypeptides and encoding polynucleotides. The HAM polypeptides can be
CC expressed by standard recombinant methodology. The HAM polypeptides are
CC useful for treating HAM-associated disorder consisting of inflammatory,
CC autoimmune, graft-versus-host, neurological, myelination, cell
CC proliferative, cardiovascular, haematologic, liver, metabolic, weight or
CC bone disorder. The present sequence represents a cDNA encoding a mouse
CC HAM polypeptide
XX
SQ Sequence 4134 BP; 1085 A; 981 C; 1036 G; 1032 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.42e-262	Length:	4134
Score:	4623.00	Matches:	848
Percent Similarity:	73.74%	Conservative:	205
Best Local Similarity:	59.38%	Mismatches:	300
Query Match:	58.41%	Indels:	75
DB:	7	Gaps:	24

US-09-787-097-12 (1-1429) x ABZ58710 (1-4134)

```
Qy         26 SerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArgProGlyLeu------ 43

Db          6 GCCGGGGGTCCG---GGCCCGCTCGGGTGCCCCGCAGCCGGCCTCCCCGGTGCTGTGGAG 62

Qy         44 ---GlyAlaGlyLeuArgLeuProArgLeuLeuSerProProLeuArgProArgLeuLeu 62

Db         63 GGCTCGGCCGGCGGGCGGTGGGGGCGCCTCCTCCTGGCTGCTGCTGGACGGGAACAGCTG 122

Qy         63 LeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeuLeuProCysGluAlaGlu 82

Db        123 GCTGCTGTGCTATGGCTTCCTCTACCT---------------------------------- 149

Qy         83 AlaAlaAlaAlaAlaAlaAlaValSerGly-SerAlaAlaAlaGluAlaLysGluCysAs 102

Db        150 ------------GGCGCTCTATGCTCAGGTGTCC---------CAGTCCAAGCCCTGCGA 188

Qy        102 pArg------ProCysValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCy 120

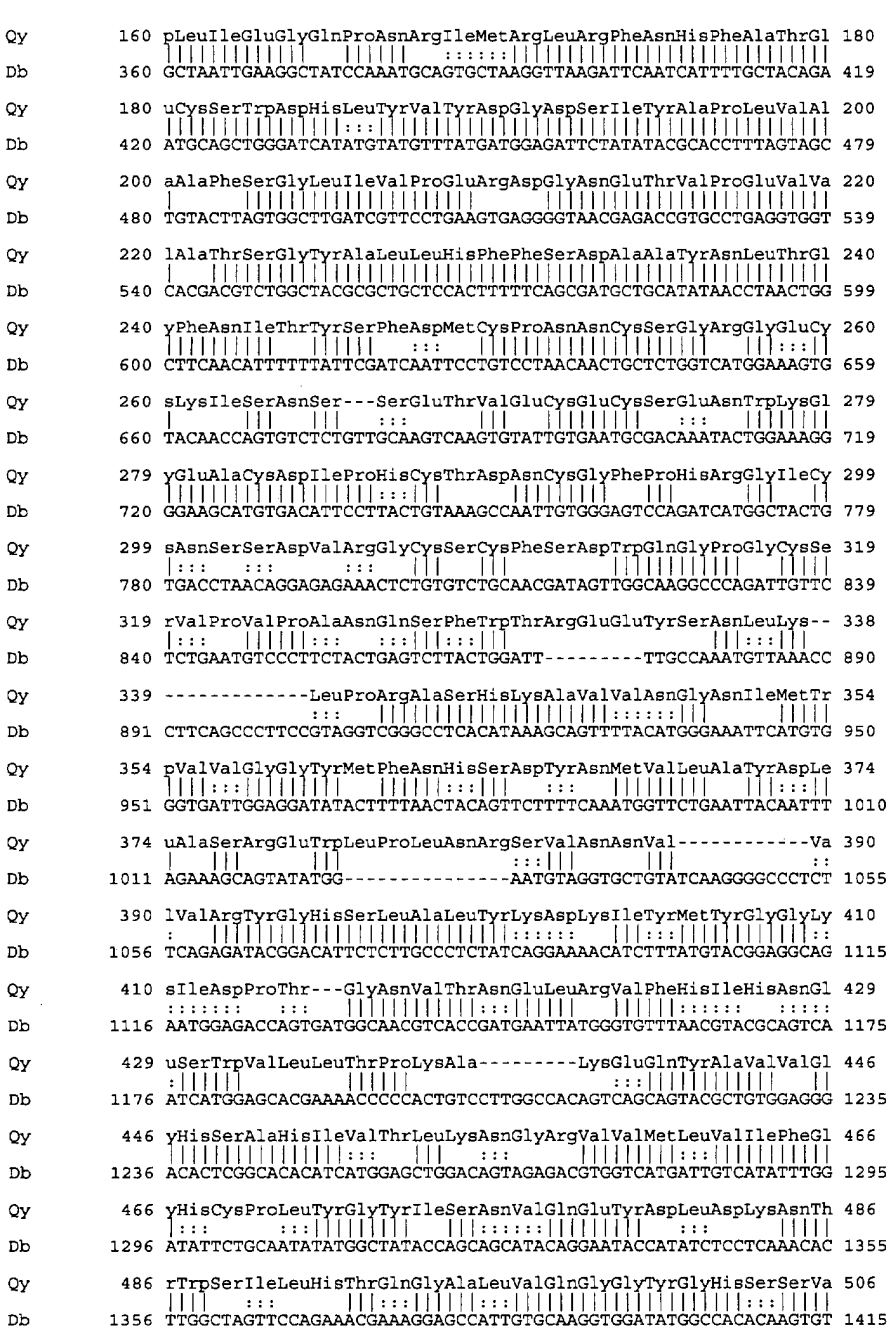
Db        189 GAGGACTGGCTCCTGCTTCTCC---GGTCGCTGTGTCAACTCCACC------TGCCTGTG 239

Qy        120 sProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGl 140

Db        240 CGACCCGGGCTGGGTTGGGGACCAGTGCCAGCACTGCCAGGGCAGGTTCAAGTTAACAGA 299

Qy        140 ySerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTr 160

Db        300 ACCTTCTGGATATTTAACAGATGGACCAATTAACTATAAATATAAAACAAAGTGTACATG 359
```


Qy 160 pLeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGl 180
Db 360 GCTAATTGAAGGCTATCCAAATGCAGTGCTAAGGTTAAGATTCAATCATTTTGCTACAGA 419
Qy 180 uCysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAl 200
Db 420 ATGCAGCTGGGATCATATGTATGTTTATGATGGAGATTCTATATACGCACCTTTAGTAGC 479
Qy 200 aAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGluValVa 220
Db 480 TGTACTTAGTGGCTTGATCGTTCCTGAAGTGAGGGGTAACGAGACCGTGCCTGAGGTGGT 539
Qy 220 lAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGl 240
Db 540 CACGACGTCTGGCTACGCGCTGCTCCACTTTTTCAGCGATGCTGCATATAACCTAACTGG 599
Qy 240 yPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCy 260
Db 600 CTTCAACATTTTTTATTCGATCAATTCCTGTCCTAACAACTGCTCTGGTCATGGAAAGTG 659
Qy 260 sLysIleSerAsnSer---SerGluThrValGluCysGluCysSerGluAsnTrpLysGl 279
Db 660 TACAACCAGTGTCTCTGTTGCAAGTCAAGTGTATTGTGAATGCGACAAATACTGGAAAGG 719
Qy 279 yGluAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCy 299
Db 720 GGAAGCATGTGACATTCCTTACTGTAAAGCCAATTGTGGGAGTCCAGATCATGGCTACTG 779
Qy 299 sAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSe 319
Db 780 TGACCTAACAGGAGAGAAACTCTGTGTCTGCAACGATAGTTGGCAAGGCCCAGATTGTTC 839
Qy 319 rValProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeuLys-- 338
Db 840 TCTGAATGTCCCTTCTACTGAGTCTTACTGGATT---------TTGCCAAATGTTAAACC 890
Qy 339 ------------LeuProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTr 354
Db 891 CTTCAGCCCTTCCGTAGGTCGGGCCTCACATAAAGCAGTTTTACATGGGAAATTCATGTG 950
Qy 354 pValValGlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLe 374
Db 951 GGTGATTGGAGGATATACTTTTAACTACAGTTCTTTTCAAATGGTTCTGAATTACAATTT 1010
Qy 374 uAlaSerArgGluTrpLeuProLeuAsnArgSerValAsnAsnVal----------Va 390
Db 1011 AGAAAGCAGTATATGG--------------AATGTAGGTGCTGTATCAAGGGGCCCTCT 1055
Qy 390 lValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLy 410
Db 1056 TCAGAGATACGGACATTCTCTTGCCCTCTATCAGGAAAACATCTTTATGTACGGAGGCAG 1115
Qy 410 sIleAspProThr---GlyAsnValThrAsnGluLeuArgValPheHisIleHisAsnGl 429
Db 1116 AATGGAGACCAGTGATGGCAACGTCACCGATGAATTATGGGTGTTTAACGTACGCAGTCA 1175
Qy 429 uSerTrpValLeuLeuThrProLysAla---------LysGluGlnTyrAlaValValGl 446
Db 1176 ATCATGGAGCACGAAAACCCCCACTGTCCTTGGCCACAGTCAGCAGTACGCTGTGGAGGG 1235
Qy 446 yHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValIlePheGl 466
Db 1236 ACACTCGGCACACATCATGGAGCTGGACAGTAGAGACGTGGTCATGATTGTCATATTTGG 1295
Qy 466 yHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsnTh 486
Db 1296 ATATTCTGCAATATATGGCTATACCAGCAGCATACAGGAATACCATATCTCCTCAAACAC 1355
Qy 486 rTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisSerServa 506
Db 1356 TTGGCTAGTTCCAGAAACGAAAGGAGCCATTGTGCAAGGTGGATATGGCCACACAAGTGT 1415
Qy 506 lTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaAs 526
Db 1416 GTATGACGAAGTGACCAAGTCCATCTATGTTCACGGAGGCTACAAAGCATTGCCTGGCAA 1475
Qy 526 nLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetTrpThrIl 546
Db 1476 TAAGTACGGGCTGGTGGACGACCTCTATAAGTACGAAGTCAACACCAGGACTTGGACTAT 1535
Qy 546 eLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerGlyThrMe 566
Db 1536 TTTGAAAGAAAGTGGGTTTGCCAGATACCTTCACTCAGCTGTTCTTATCAATGGAGCTAT 1595
Qy 566 tLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysPh 586
Db 1596 GCTTATTTTTGGAGGAAATACCCATAATGATACTTCCCTGAGTAACGGTGCAAAATGTTT 1655
Qy 586 eSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuProArgPr 606
Db 1656 TTCTGCCGATTTCCTGGCATATGACATAGCTTGTGACGAATGGAAGACATTGCCTAAACC 1715
Qy 606 oAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrMe 626
Db 1716 AAATCTCCATAGAGACGTCAACCGATTTGGGCATTCTGCAGTCGTCATCAATGGGTCAAT 1775
Qy 626 tTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSerGl 646
Db 1776 GTATATATTTGGTGGATTTTCTAGTGTACTCCTTAATGATATCCTTGTGTATAAACCCCC 1835
Qy 646 uGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgCy 666
Db 1836 AAATTGCAAAGCTTTCCGAGATGAAGAACTGTGCAGAAACGCTGGTCCAGGGATAAAATG 1895
Qy 666 sValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnGl 686
Db 1896 TGTTTGGAAT------AAGAATCACTGTGAATCTTGG------GAGTCTGGGAATACAAA 1943
Qy 686 uGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArgCysAspGl 706
Db 1944 TAATATTCTCAGAGCCAAGTGCCCTCCCAAGACAGCTGCTACCGATGACAGATGTTACAG 2003
Qy 706 nHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAsp-- 725
Db 2004 ATATGCTGACTGTGCCAGCTGCACAGCCAACACGAACGGGTGCCAGTGGTGTGACGACAA 2063
Qy 726 -HisCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGl 745
Db 2064 GAAATGCATCTCAGCCAGCAGCAACTGCAGC---------ACGTCTGTCAGAAACTACAC 2114
Qy 745 uAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArgSerCy 765
Db 2115 TAAATGTCATATAAGAAATGAGCAG---ATTTGTAACAAACTTACAAGCTGTAAAAGCTG 2171
Qy 765 sAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuProGl 785
Db 2172 TTCACTCAACTTGAATTGCCAGTGGGATCAGCGGCAGCAGGAATGTCAGGCTTTGCCAGC 2231
Qy 785 uAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThrThrAl 805
Db 2232 TCACCTTTGTGGAGAAGGCTGGAATCATGTTGGGGACGCTTGTCTTCGAATCAATTCCAG 2291
Qy 805 aLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSe 825
Db 2292 TCGAGAAAGCTATGATAATGCCAAACTTTATTGCTATAACCTCAGTGGAAATCTCGCCTC 2351
Qy 825 rLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGlnSerSe 845
Db 2352 CCTGACCACGTCCAAGGAGGTGGAGTTTGTGTTGGATGAAATA---------- 2394
Qy 845 rGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSe 865
Db 2395 -CAGAAGTTCACACAGCAGAAAGTGTCACCGTGGGTAGGCCTACGCAAGATCAACATATC 2453
Qy 865 rTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSe 885

```
Db       2454 CTACTGGGGATGGGAGGACATGTCTCCTTTCACAAATACAAGCCTGCAGTGGCTTCCTGG 2513

Qy        885 rGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLy 905

Db       2514 TGAGCCAAATGACTCTGGATTCTGTGCCTACTTAGAAAGGGCTGCAGTGGCAGGGTTAAA 2573

Qy        905 sAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAla-------- 922

Db       2574 AGCAAACCCTTGCACATCCATGGCAGATGGACTCGTTTGTGAAAAGCCTGTAGTAAGCCC 2633

Qy        923 -AsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCy 942

Db       2634 AAATCAGAACGCGAGGCCGTGCAAGAAGCCGTGCTCCCTGAGGACCTCCTGCGCCAACTG 2693

Qy        942 sThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAs 962

Db       2694 CACGAGCAGCGGCATGGAGTGCATGTGGTGCAGCAGCACGAAGCGCTGTGTGGACTCCAA 2753

Qy        962 nAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSerThrCy 982

Db       2754 CGCTTACATCATCTCCTTTCCCTACGGACAGTGCCTGGAGTGGCAGACT---GCCACCTG 2810

Qy        982 sProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCy 1002

Db       2811 CTCGCCTCAAAATTGTTCTGGGTTAAGAACCTGTGGACAGTGCTTGGAGCAGCCAGGGTG 2870

Qy       1002 sGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLy 1022

Db       2871 TGGCTGGTGCAACGATCCTAGTAACACAGGAAGAGGCTATTGCATCGAAGGGTCTTCCCG 2930

Qy       1022 sGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLe 1042

Db       2931 GGGCCCAATGAAACTC--------------GTGGGGGTCCACAACAGTGACGTGGTTCT 2975

Qy       1042 uAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAl 1062

Db       2976 TGACACCAGCCTCTGCCCCAAGGAGAAGAACTACGAGTGGTCTTTTATCCAGTGTCCAGC 3035

Qy       1062 aCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAs 1082

Db       3036 TTGCCAGTGTAATGGACACAGCACGTGCATCAACAACAACGTCTGTGAGCAGTGTAAGAA 3095

Qy       1082 nLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAs 1102

Db       3096 TCTCACCACTGGGCGACAGTGTCAGGAATGCATGCCAGGGTACTATGGAGACCCAACCAA 3155

Qy       1102 nGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnTh 1122

Db       3156 CGGTGGGCAGTGCACAGCTTGCACGTGCGGCGGCCATGCGAACGTCTGTCACCTGCACAC 3215

Qy       1122 rGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluVa 1142

Db       3216 GGGAAAGTGTTTTTGCACAACCAAGGGGATCAAGGGTGACCAGTGCCAGCTATGTGACTC 3275

Qy       1142 lGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAs 1162

Db       3276 TGAAAATCGCTATGTTGGTAATCCACTTAGGGGGACATGCTATTACAGTCTTCTGATTGA 3335

Qy       1162 pTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIleAsnPh 1182

Db       3336 CTACCAGTTTACCTTCAGCTTGCTGCAGGAAGATGACCGGCACCACACTGCCATCAACTT 3395

Qy       1182 eValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAs 1202

Db       3396 CATCGCCAACCCAGAGCAGTCAAACAAAAACTTGGACATTTCGATTAATGCTTCCAACAA 3455

Qy       1202 nPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGl 1222

Db       3456 CTTTAATCTCAACATTACGTGGTCAGTTGGCTCAACAGGTGGAACCATATCTGGGGAAGA 3515

Qy       1222 uMetProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLy 1242

Db       3516 GACTCCTATAGTTTCTAAGACAAATATAAAGGAATACAGAGACAGCTTTTCCTATGAAAA 3575

Qy       1242 sPheAspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTr 1262

Db       3576 ATTTAACTTCAGAAGCAATCCTAATATCACATTTTATGTGTATGTCAGCAACTTCTCCTG 3635

Qy       1262 pProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPh 1282

Db       3636 GCCTATTAAAATACAGATTGCGTTTTCACAACACAACACGATCATGGATCTCGTGCAGTT 3695

Qy       1282 ePheValThrPhePheSerCysPheLeuSerLeuLeuLeuValAlaAlaValValTrpLy 1302

Db       3696 CTTTGTCACCTTCTTCAGTTGTTTTTTATCTTTACTGCTGGTGGCTGCTGTGGTCTGGAA 3755

Qy       1302 sIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeuLeuArgGluMetGlnGl 1322

Db       3756 GATCAAACAAACTTGCTGGGCTTCTCGTCGGAGGGAGCAACTGCTTCGAGAACGACAGCA 3815

Qy       1322 nMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThrAspGluGluProPr 1342

Db       3816 GATGGCCAGCCGTCCCTTTGCTTCTGTTGATGTAGCCCTGGAAGTAGGAGCTGAACAGAC 3875

Qy       1342 oAspLeuIleGlyGlySerIleLysThrValProLysProIleAlaLeuGluProCysPh 1362

Db       3876 AGACTTTCTGCGAGGGCCATTAGAGGGTGCCCCTAAGCCAATAGCCATCGAACCCTGCGC 3935

Qy       1362 eGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyGlyIl 1382

Db       3936 TGGGAACAGAGCTGCTGTCCTGACTGTGTTTCTCTGTCTACCGAGAGGATCTTCAGGCGC 3995

Qy       1382 eProProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGl 1402

Db       3996 CCCACCCCCTGGGCAGTCAGGCCTTGCTATCGCCAGTGCCCTGATAGACATCTCACAGCA 4055

Qy       1402 nMetProIleValTyrLysGluLysSerGlyAlaValArgAsnArgLysGlnGlnProPr 1422

Db       4056 GAAGCCTTCTGATAATAAAGACAAGACTTCTGGAGTCCGCAATCGGAAGCACCTCTCCAC 4115

Qy       1422 oAlaGlnProGlyThrCysIle 1429

Db       4116 ACGTCAA---GGAACTTGTGTC 4134
```

RESULT 15
AAZ91922
ID AAZ91922 standard; cDNA; 2625 BP.
XX
AC AAZ91922;
XX
DT 08-JUN-2000 (first entry)
XX
DE Human mahogany protein coding sequence #4.
XX
KW Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;
KW weight regulation; cell therapy; body weight disorder; cachexia;
KW anorexia; hyperpigmentation; increased metabolic rate disorder;
KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.
XX
OS Homo sapiens.
XX
PN WO200005373-A2.
XX
PD 03-FEB-2000.
XX
PF 21-JUL-1999; 99WO-US016484.
XX
PR 21-JUL-1998; 98US-0093630P.
PR 20-OCT-1998; 98US-0104978P.
PR 05-FEB-1999; 99US-00245041.
XX
PA (MILL-) MILLENIUM PHARM INC.
XX
PI Moore K, Nagle DL;
XX

```
DR   WPI; 2000-195103/17.
DR   P-PSDB; AAY81809.
XX
PT   New human and murine mahogany genes, useful, e.g. for diagnosis and
PT   treatment of body weight disorders.
XX
PS   Claim 1; Fig 20a; 188pp; English.
XX
CC   This sequence represents a human mahogany gene of the invention. The
CC   mahogany genes are used: (i) to produce recombinant mahogany (mg)
CC   proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming
CC   therapeutics; (iii) as a source of diagnostic probes and primers for
CC   detecting expression of mg genes or mutations, regulatory defects, in
CC   this gene, or for isolation of related sequences; and (iv) in (cell-
CC   based) gene therapy. (II) are used to raise specific antibodies (Ab); to
CC   identify other (extra)cellular products involved in weight regulation,
CC   and to screen for agents that disrupt interaction between (II) and other
CC   macromolecules. The Ab are used to detect abnormal levels (or function)
CC   of (II) (for diagnosis, prognosis or monitoring of treatment); to
CC   evaluate (II)-expressing cells intended for cell therapy, and as
CC   therapeutic mg inhibitors. Cells that express the mg gene (or contain the
CC   mg polypeptide) are used to identify agents (A) that modulate mg
CC   activity. (A) are potentially useful for the treatment of body weight
CC   disorders, particularly obesity, cachexia or anorexia, or other
CC   conditions associated with the mg gene such as hyperpigmentation,
CC   hyperphagia and disorders that result in increased metabolic rate
XX
SQ   Sequence 2625 BP; 660 A; 605 C; 694 G; 666 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                3.64e-234      Length:          2625
Score:                    4136.50        Matches:         759
Percent Similarity:       95.48%         Conservative:    2
Best Local Similarity:    95.23%         Mismatches:      22
Query Match:              52.26%         Indels:          14
DB:                       3              Gaps:            3

US-09-787-097-12 (1-1429) x AAZ91922 (1-2625)

Qy          2 ValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAlaAla 21
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         10 GTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCG 69

Qy         22 LeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArgPro 41
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 CTCGCGGGCAGGAGCGGCGGGCCGCACCGACCCTGCACCGCGACAGGGGCCTGGAGGCCG 129

Qy         42 GlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSer---------------ProPro 56
              |||  |||  |||  |||||||||:::||||||                  ||||||
Db        130 GGACCGCGCGCCCGGCTGTGTCTCCCGCGGGTGCTGTCGCGGGCGCTGCCCCCGCCGCCG 189

Qy         57 LeuArgProArgLeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeu 76
              |||  |||  ||||||      ||||||||||||||||
Db        190 CTGCTGCCG---CTGCTCTTTTCGCTGCTGCTGCTGCCG---------------------- 225

Qy         77 LeuProCysGluAlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAla 96
              ||||||  ||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db        226 CTGCCCCGGGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCC 285

Qy         97 GluAlaLysGluCysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGly 116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 GAGGCCAAGGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGC 345

Qy        117 GlnCysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPhe 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 CAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTC 405

Qy        137 ArgLeuThrGlySerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThr 156
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        406 AGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACG 465

Qy        157 LysCysThrTrpLeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHis 176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        466 AAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCAT 525

Qy        177 PheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAla 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        526 TTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCA 585

Qy        197 ProLeuValAlaAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrVal 216
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        586 CCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTC 645

Qy        217 ProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyr 236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        646 CCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTAT 705

Qy        237 AsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGly 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        706 AATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGC 765

Qy        257 ArgGlyGluCysLysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsn 276
              ||||||||||||||||||||||||||||||:::||||||||||||||||||||||||||
Db        766 CGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAAC 825

Qy        277 TrpLysGlyGluAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArg 296
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        826 TGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGA 885

Qy        297 GlyIleCysAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyPro 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        886 GGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCT 945

Qy        317 GlyCysSerValProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsn 336
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        946 GGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAAC 1005

Qy        337 LeuLysLeuProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValVal 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1006 TTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTT 1065

Qy        357 GlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSer 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1066 GGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCT 1125

Qy        377 ArgGluTrpLeuProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSer 396
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1126 AGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCT 1185

Qy        397 LeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsn 416
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1186 TTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAAT 1245

Qy        417 ValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrPro 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1246 GTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCT 1305

Qy        437 LysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsn 456
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1306 AAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAAT 1365

Qy        457 GlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsn 476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1366 GGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAAT 1425

Qy        477 ValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeu 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1426 GTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTT 1485

Qy        497 ValGlnGlyGlyTyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrVal 516
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1486 GTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTT 1545

Qy        517 HisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArg 536
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1546 CATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGA 1605
```

```
Qy        537 TyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeu 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1606 TATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTG 1665

Qy        557 HisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAsp 576
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1666 CACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGAC 1725

Qy        577 ThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAla 596
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1726 ACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCC 1785

Qy        597 CysAspArgTrpSerValLeuProArgProAspLeuHisHisAspValAsnArgPheGly 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1786 TGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGC 1845

Qy        617 HisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeu 636
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1846 CATTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTC 1905

Qy        637 LeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAla 656
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1906 CTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCT 1965

Qy        657 CysLeuAlaAlaGlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIle 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1966 TGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATC 2025

Qy        677 SerTrpAlaLeuAlaThrAspGluGlnGluGluLysLeuLysSerGluCysPheSerLys 696
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2026 TCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAA 2085

Qy        697 ArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsn 716
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2086 AGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAAC 2145

Qy        717 ThrAsnAspCysHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGlu 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2146 ACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAA 2205

Qy        737 GlyGlnIleSerIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCys 756
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2206 GGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGT 2265

Qy        757 AsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArg 776
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2266 AACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGG 2325

Qy        777 AsnGlnGluCysIleAlaLeuProGluAsnIleCys-GlyIleGlyTrp 792
              ||||||||||||||||||||||          ||| |||     |||
Db       2326 AATCAGGAGTGCATTGCCCTGCCCGGTAGGCCTTGCAGGGTCATCTTGG 2374
```

Search completed: March 1, 2004, 02:00:36
Job time : 1389 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 1, 2004, 01:57:50 ; Search time 1024 Seconds
(without alignments)
5034.248 Million cell updates/sec

Title: US-09-787-097-12
Perfect score: 7915
Sequence: 1 MVAAAAATEARLRRRTAATA.....KSGAVRNRKQQPPAQPGTCI 1429

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2353733 seqs, 1803733377 residues

Total number of hits satisfying chosen parameters: 4707466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09787097/runat_26022004_062937_23538/app_query.fasta_1.1607
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09787097_@CGN_1_1_136_@runat_26022004_062937_23538
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	7696.5	97.2	8589	9	US-09-893-238-14	Sequence 14, Appl
	2	7303	92.3	8827	9	US-09-893-238-1	Sequence 1, Appli
	3	6870.5	86.8	4072	9	US-09-893-238-16	Sequence 16, Appl
	4	4637	58.6	6733	16	US-10-197-824-1	Sequence 1, Appli
	5	4136.5	52.3	2625	9	US-09-893-238-18	Sequence 18, Appl
	6	3491.5	44.1	6370	9	US-09-893-238-12	Sequence 12, Appl
	7	3435	43.4	2419	9	US-09-893-238-8	Sequence 8, Appli
	8	1226.5	15.5	1051	9	US-09-893-238-10	Sequence 10, Appl
	9	960.5	12.1	8589	15	US-10-085-198-15	Sequence 15, Appl
	10	923	11.7	7244	15	US-10-120-988-143	Sequence 143, App
	11	840	10.6	207433	15	US-10-277-216-5	Sequence 5, Appli
	12	840	10.6	207433	16	US-10-126-022-5	Sequence 5, Appli
	13	766	9.7	90050	9	US-09-893-238-5	Sequence 5, Appli
	14	581.5	7.3	415	9	US-09-833-381-1916	Sequence 1916, Ap
c	15	536	6.8	625	14	US-10-198-846-5869	Sequence 5869, Ap
c	16	454	5.7	17056	9	US-09-893-238-3	Sequence 3, Appli
c	17	446	5.6	246	9	US-09-736-457-966	Sequence 966, App
c	18	446	5.6	246	9	US-09-902-941-966	Sequence 966, App
c	19	446	5.6	246	9	US-09-849-626-966	Sequence 966, App
c	20	446	5.6	246	14	US-10-017-754-966	Sequence 966, App
c	21	446	5.6	246	14	US-10-113-872-966	Sequence 966, App
	22	431	5.4	531	9	US-09-864-761-15900	Sequence 15900, A
	23	393	5.0	212	9	US-09-796-692-7610	Sequence 7610, Ap
	24	393	5.0	212	14	US-10-040-862-7610	Sequence 7610, Ap
	25	393	5.0	212	15	US-10-057-475B-7610	Sequence 7610, Ap
	26	393	5.0	212	15	US-10-154-884B-7610	Sequence 7610, Ap
	27	369.5	4.7	8064	14	US-10-004-113-56	Sequence 56, Appl
	28	365	4.6	7596	14	US-10-004-113-57	Sequence 57, Appl
	29	359	4.5	7693	12	US-10-072-012-133	Sequence 133, App
	30	357	4.5	7332	9	US-09-944-849-1	Sequence 1, Appli
	31	357	4.5	7673	9	US-09-815-925-1	Sequence 1, Appli
	32	357	4.5	7673	15	US-10-294-006-1	Sequence 1, Appli
	33	356.5	4.5	6728	15	US-10-190-115-3	Sequence 3, Appli
	34	356.5	4.5	6728	15	US-10-369-072-3	Sequence 3, Appli
	35	340	4.3	193	9	US-09-864-761-32404	Sequence 32404, A
	36	340	4.3	7410	15	US-10-190-115-1	Sequence 1, Appli
	37	340	4.3	7410	15	US-10-369-072-1	Sequence 1, Appli
	38	334.5	4.2	7615	14	US-10-004-113-59	Sequence 59, Appl
	39	334	4.2	5312	14	US-10-311-455-169	Sequence 169, App
c	40	332	4.2	182	9	US-09-864-761-21460	Sequence 21460, A
	41	326.5	4.1	8091	14	US-10-101-510-86	Sequence 86, Appl
	42	326.5	4.1	8091	14	US-10-356-625-1	Sequence 1, Appli
	43	326.5	4.1	8091	15	US-10-159-563-322	Sequence 322, App
c	44	326	4.1	375	9	US-09-864-761-4720	Sequence 4720, Ap
	45	303	3.8	11558	10	US-09-949-029-133	Sequence 133, App

ALIGNMENTS

RESULT 1
US-09-893-238-14
; Sequence 14, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 8589
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-893-238-14

Alignment Scores:

Pred. No.:	0	Length:	8589
Score:	7696.50	Matches:	1401
Percent Similarity:	97.91%	Conservative:	2
Best Local Similarity:	97.77%	Mismatches:	17
Query Match:	97.24%	Indels:	13
DB:	9	Gaps:	3

US-09-787-097-12 (1-1429) x US-09-893-238-14 (1-8589)

```
Qy          2 ValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAlaAla 21
Db         10 GTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCG 69

Qy         22 LeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArgPro 41
Db         70 CTCGCGGGCAGGAGCGGCGGGCCGCACCGACCCTGCACCGCGACAGGGGCCTGGAGGCCG 129

Qy         42 GlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSer----------------ProPro 56
Db        130 GGACCGCGCGCCCGGCTGTGTCTCCCGCGGGTGCTGTCGCGGGCGCTGCCCCCGCCGCCG 189

Qy         57 LeuArgProArgLeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeu 76
Db        190 CTGCTGCCG---CTGCTCTTTTCGCTGCTGCTGCTGCCG---------------------- 225

Qy         77 LeuProCysGluAlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAla 96
Db        226 CTGCCCCGGGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCC 285

Qy         97 GluAlaLysGluCysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGly 116
Db        286 GAGGCCAAGGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGC 345

Qy        117 GlnCysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPhe 136
Db        346 CAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTC 405

Qy        137 ArgLeuThrGlySerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThr 156
Db        406 AGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACG 465

Qy        157 LysCysThrTrpLeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHis 176
Db        466 AAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCAT 525

Qy        177 PheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAla 196
Db        526 TTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCA 585

Qy        197 ProLeuValAlaAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrVal 216
Db        586 CCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTC 645

Qy        217 ProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyr 236
Db        646 CCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTAT 705

Qy        237 AsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGly 256
Db        706 AATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGC 765

Qy        257 ArgGlyGluCysLysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsn 276
Db        766 CGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAAC 825

Qy        277 TrpLysGlyGluAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArg 296
Db        826 TGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGA 885

Qy        297 GlyIleCysAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyPro 316
Db        886 GGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCT 945

Qy        317 GlyCysSerValProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsn 336
Db        946 GGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAAC 1005

Qy        337 LeuLysLeuProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValVal 356
Db       1006 TTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTT 1065

Qy        357 GlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSer 376
Db       1066 GGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCT 1125

Qy        377 ArgGluTrpLeuProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSer 396
Db       1126 AGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCT 1185

Qy        397 LeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsn 416
Db       1186 TTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAAT 1245

Qy        417 ValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrPro 436
Db       1246 GTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCT 1305

Qy        437 LysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsn 456
Db       1306 AAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAAT 1365

Qy        457 GlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsn 476
Db       1366 GGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAAT 1425

Qy        477 ValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeu 496
Db       1426 GTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTT 1485

Qy        497 ValGlnGlyGlyTyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrVal 516
Db       1486 GTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTT 1545

Qy        517 HisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArg 536
Db       1546 CATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGA 1605

Qy        537 TyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeu 556
Db       1606 TATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTG 1665

Qy        557 HisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAsp 576
Db       1666 CACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGAC 1725

Qy        577 ThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAla 596
Db       1726 ACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCC 1785

Qy        597 CysAspArgTrpSerValLeuProArgProAspLeuHisHisAspValAsnArgPheGly 616
Db       1786 TGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGC 1845

Qy        617 HisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeu 636
Db       1846 CATTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTC 1905

Qy        637 LeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAla 656
Db       1906 CTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCT 1965

Qy        657 CysLeuAlaAlaGlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIle 676
```

```
Db      1966 TGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATC 2025
Qy       677 SerTrpAlaLeuAlaThrAspGluGlnGluGluLysLeuLysSerGluCysPheSerLys 696
Db      2026 TCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAA 2085
Qy       697 ArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsn 716
Db      2086 AGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAAC 2145
Qy       717 ThrAsnAspCysHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGlu 736
Db      2146 ACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAA 2205
Qy       737 GlyGlnIleSerIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCys 756
Db      2206 GGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGT 2265
Qy       757 AsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArg 776
Db      2266 AACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGG 2325
Qy       777 AsnGlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGly 796
Db      2326 AATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGA 2385
Qy       797 AsnSerCysLeuLysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCys 816
Db      2386 AACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGT 2445
Qy       817 ArgAsnHisAsnAlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeu 836
Db      2446 AGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTT 2505
Qy       837 LysGlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrp 856
Db      2506 AAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGG 2565
Qy       857 ValGlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThr 876
Db      2566 GTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACA 2625
Qy       877 AsnSerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeu 896
Db      2626 AATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTA 2685
Qy       897 SerGluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySer 916
Db      2686 TCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGT 2745
Qy       917 ValCysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArg 936
Db      2746 GTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGG 2805
Qy       937 ThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLys 956
Db      2806 ACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAG 2865
Qy       957 GlnCysValAspSerAsnAlaTyrValAlaSerProProPheGlyGlnCysMetGluTrp 976
Db      2866 CAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGG 2925
Qy       977 TyrThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCys 996
Db      2926 TATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGC 2985
Qy       997 LeuGluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCys 1016
Db      2986 TTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGC 3045
Qy      1017 IleGluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPhe 1036
Db      3046 ATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTC 3105

Qy      1037 TyrProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSer 1056
Db      3106 TATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCT 3165
Qy      1057 PheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIle 1076
Db      3166 TTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATC 3225
Qy      1077 CysGluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPhe 1096
Db      3226 TGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTC 3285
Qy      1097 TyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSer 1116
Db      3286 TACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCT 3345
Qy      1117 LeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGlu 1136
Db      3346 CTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAG 3405
Qy      1137 CysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyr 1156
Db      3406 TGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTAT 3465
Qy      1157 TyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyr 1176
Db      3466 TATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTAT 3525
Qy      1177 TyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPhe 1196
Db      3526 TACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTC 3585
Qy      1197 IleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGly 1216
Db      3586 ATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGA 3645
Qy      1217 ThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAsp 1236
Db      3646 ACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGAT 3705
Qy      1237 SerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyr 1256
Db      3706 AGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTAT 3765
Qy      1257 ValSerAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPhe 1276
Db      3766 GTCAGTAATTTCACCTGGCCCATCAAAATTCAGATTGCCTTCTCTCAGCACAGCAATTTT 3825
Qy      1277 MetAspLeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuLeuVal 1296
Db      3826 ATGGACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTG 3885
Qy      1297 AlaAlaValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeu 1316
Db      3886 GCTGCTGTGGTTTGGAAGATCAAACAAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTT 3945
Qy      1317 LeuArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGlu 1336
Db      3946 CTTCGAGAGATGCAACAGATGGCCAGCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAA 4005
Qy      1337 ThrAspGluGluProProAspLeuIleGlyGlySerIleLysThrValProLysProIle 1356
Db      4006 ACAGATGAGGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATT 4065
Qy      1357 AlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuPro 1376
Db      4066 GCACTGGAGCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCT 4125
Qy      1377 ArgGlyLeuGlyGlyIleProProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeu 1396
Db      4126 CGAGGCCTGGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTG 4185
```

```
Qy       1397 ValAspIleSerGlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAsn 1416
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4186 GTGGACATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAAC 4245

Qy       1417 ArgLysGlnGlnProProAlaGlnProGlyThrCysIle 1429
             |||||||||||||||||||||||||||||||||||||
Db       4246 CGGAAGCAGCAGCCCCCTGCACAGCCTGGGACCTGCATC 4284

RESULT 2
US-09-893-238-1
; Sequence 1, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT:  Nagle, D.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
;  FILE REFERENCE: 7853-237
;  CURRENT APPLICATION NUMBER: US/09/893,238
;  CURRENT FILING DATE:  2001-06-27
;  PRIOR APPLICATION NUMBER: 09/245,041
;  PRIOR FILING DATE: 1999-02-05
;  PRIOR APPLICATION NUMBER: 60/093,630
;  PRIOR FILING DATE: 1998-07-21
;  PRIOR APPLICATION NUMBER: 60/104,978
;  PRIOR FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 8827
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-893-238-1

Alignment Scores:
Pred. No.:              0           Length:         8827
Score:                  7303.00     Matches:        1332
Percent Similarity:     90.99%      Conservative:   32
Best Local Similarity:  88.86%      Mismatches:     56
Query Match:            92.27%      Indels:         79
DB:                     9           Gaps:           4

US-09-787-097-12 (1-1429) x US-09-893-238-1 (1-8827)

Qy          2 ValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAlaAla 21
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 GTGGCGGCGGCGGCGGCGACTGAGGCGCGGCTGAGGGGGAGCACGAGGACGACAGCAGCG 142

Qy         22 LeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArgPro 41
              |||||||||  |||     |||                 |||  |||  ||||||
Db        143 CCTGCGGGCAGGAAGGGCAGGCAGCACCGACCCTGCACCGCGACAGGGGCCTGGAGGCCG 202

Qy         42 GlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSer---------------ProPro 56
              |||     |||  |||  |||||||||:::||||||                 ||||||
Db        203 GGACCGCGCGCCCGGCTGTGTCTCCCGCGGGTGCTGTCGCGGGCGCTGCCCCCGCCGCCG 262

Qy         57 LeuArgProArgLeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeu 76
              |||  |||  ||||||      ||||||||||||||||
Db        263 CTGCTGCCG---CTGCTCTTTTCGCTGCTGCTGCTGCCG---------------------- 298

Qy         77 LeuProCysGluAlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAla 96
              ||||||  ||||||||||||||||||  |||||||||||||||||||||||||||||||
Db        299 CTGCCCCGGGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCC 358

Qy         97 GluAlaLysGluCysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGly 116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 GAGGCCAAGGAATGTGACCGGCCGTGTGTCAACGGCGGCCGCTGCAACCCTGGCACCGGC 418

Qy        117 GlnCysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPhe 136
              ||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db        419 CAGTGCGTCTGCCCCACGGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTC 478

Qy        136 ------------------------------------------------------------ 136

Db        479 AGGACATCTGTCTCACGCCTATAATCACAGCTGTTCGGAAGGTGAGGCTGGAGGAACAGT 538

Qy        136 ------------------------------------------------------------ 136

Db        539 TCGAGGCAAGCTTCGGCTACAGAATAAGTTCAAGAGTAACCTGGGGCAACTTGGGCTTGT 598

Qy        136 ------------------------------------------------------------ 136

Db        599 CTCCAAAACCAAAATGAGCGAAAAGGAGCAAGCTAGAGTCTTTTGGGAAAATTTTAGCTG 658

Qy        137 --------------Arg-LeuThrGlySerSerGlyPheValThrAspGlyProGlyAs 151
                           ||| ||||||||||||||||||||||||||||||||||||||||
Db        659 ACTAATTTTTCACCGAGAACTAACTGGCTCTTCTGGATTTGTAACAGATGGACCTGGGAA 718

Qy        151 nTyrLysTyrLysThrLysCysThrTrpLeuIleGluGlyGlnProAsnArgIleMetAr 171
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 TTATAAATATAAGACGAAGTGCACATGGCTCATTGAAGGACAGCCAAATAGAATAATGAG 778

Qy        171 gLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGl 191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 ACTTCGCTTCAACCATTTTGCTACAGAATGTAGCTGGGACCATTTATATGTTTATGATGG 838

Qy        191 yAspSerIleTyrAlaProLeuValAlaAlaPheSerGlyLeuIleValProGluArgAs 211
              ||||||||||||||||||||||:::||||||||||||||||||||||||||||||||||
Db        839 GGACTCAATCTACGCACCTCTGATTGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGA 898

Qy        211 pGlyAsnGluThrValProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePh 231
              |||||||||||||   |||||||||      ||||||||||||||||||||||||||||
Db        899 TGGCAATGAGACGGCTCCTGAGGTCACTGTCACTTCAGGTTATGCACTGCTGCATTTTTT 958

Qy        231 eSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysPr 251
              ||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db        959 CAGTGATGCTGCTTATAATCTGACTGGATTTAATATCACTTACAATTTTGACATGTGTCC 1018

Qy        251 oAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSerSerGluThrValGluCy 271
              |||||||||||||||||||||||||||||||||   ||||||||||||      ||||||||
Db       1019 GAATAATTGCTCAGGCCGAGGAGAGTGTAAGAGCAGTAACAGCAGCAGCGCTGTTGAGTG 1078

Qy        271 sGluCysSerGluAsnTrpLysGlyGluAlaCysAspIleProHisCysThrAspAsnCy 291
              |||||||||||||||||||||||||||||:::||||||||||||||||||||||||||||
Db       1079 TGAATGTTCTGAAAACTGGAAAGGGGAGTCGTGTGACATTCCTCACTGTACAGACAACTG 1138

Qy        291 sGlyPheProHisArgGlyIleCysAsnSerSerAspValArgGlyCysSerCysPheSe 311
              |||||||||||||||||||||||||||:::||||||    ||||||||||||||||||||
Db       1139 TGGCTTTCCTCACCGAGGCATCTGTAATGCAAGCGATACCAGAGGGTGCTCCTGCTTTCC 1198

Qy        311 rAspTrpGlnGlyProGlyCysSerValProValProAlaAsnGlnSerPheTrpThrAr 331
                  ||||||||||||||||||||:::||||||||||||||||||||||||||||||||||
Db       1199 TCACTGGCAGGGTCCTGGATGTTCAATTCCTGTGCCAGCTAACCAGTCTTTTTGGACTCG 1258

Qy        331 gGluGluTyrSerAsnLeuLysLeuProArgAlaSerHisLysAlaValValAsnGlyAs 351
              ||||||||||||:::||||||||||||||||||||||||||||||||||||||||||||
Db       1259 AGAAGAATATTCTGATTTAAAGCTTCCCAGAGCCTCTCATAAAGCTGTGGTCAATGGAAA 1318

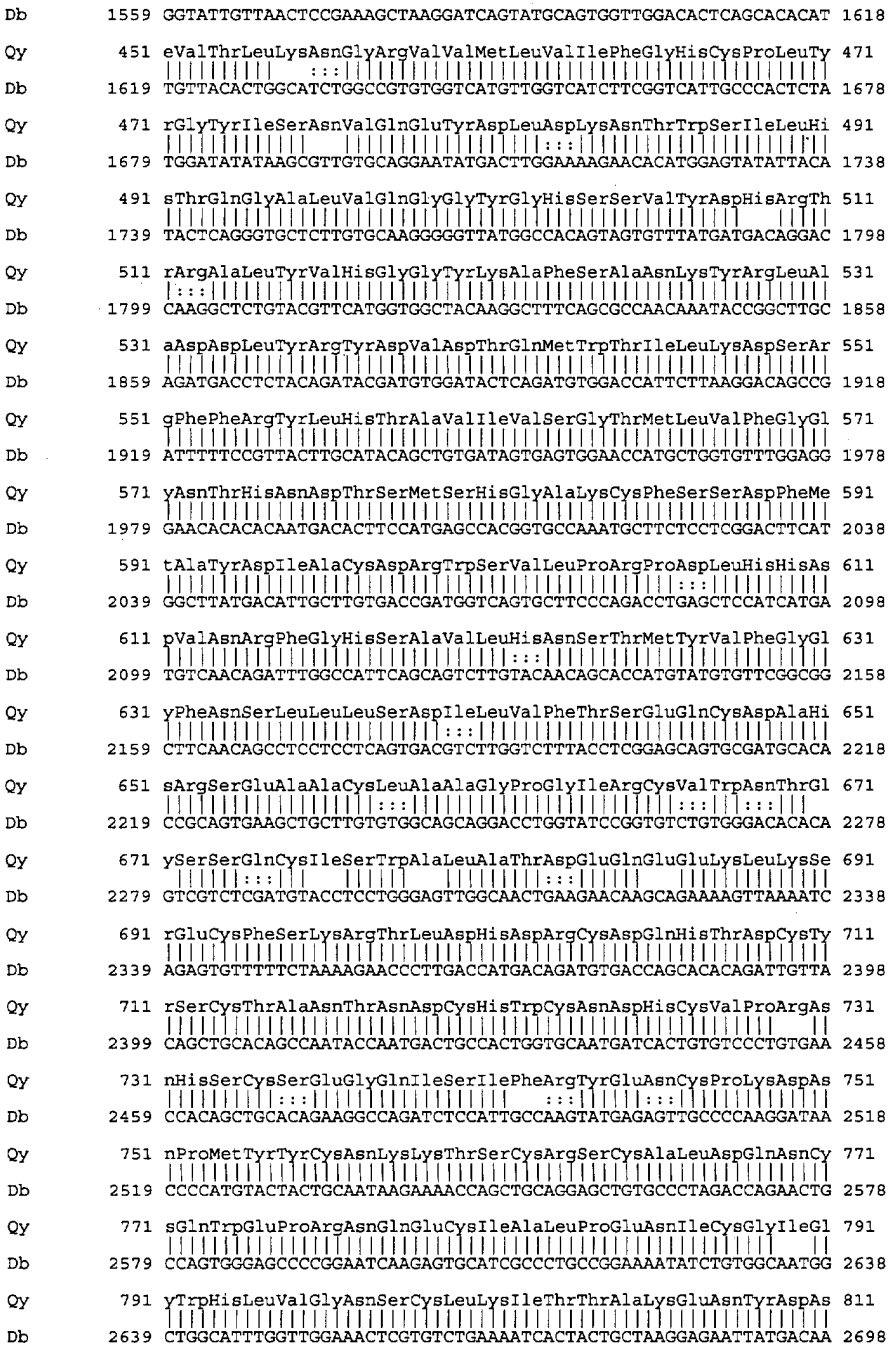
Qy        351 nIleMetTrpValValGlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAl 371
              ||||||||||||||||||||||||||||||||||||||||||||||||||:::||||||||
Db       1319 TATAATGTGGGTTGTTGGCGGATATATGTTCAACCATTCAGATTACAGCATGGTTTTAGC 1378

Qy        371 aTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSerValAsnAsnValValVa 391
              |||||||||  ||||||||||||||||||||||||  |||||||||:::||||||||
Db       1379 GTATGACCTGACTTCTAGGGAATGGCTTCCACTAAACCATTCTGTGAACAGTGTGGTTGT 1438

Qy        391 lArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLysIl 411
              ||||||||||||||||||||||:::||||||||||||||||||||||||||||||||||||
Db       1439 AAGATATGGTCATTCTTTGGCATTACATAAGGATAAAATCTACATGTATGGAGGAAAAAT 1498

Qy        411 eAspProThrGlyAsnValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTr 431
              ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1499 TGATTCAACAGGGAACGTGACCAATGAGCTGAGAGTATTTCATATTCATAATGAATCATG 1558

Qy        431 pValLeuLeuThrProLysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIl 451
              ||||||||||||||||||||||||:::||||||||||||||||||||||||||||||||||
```


Db 1559 GGTATTGTTAACTCCGAAAGCTAAGGATCAGTATGCAGTGGTTGGACACTCAGCACACAT 1618
Qy 451 eValThrLeuLysAsnGlyArgValValMetLeuValIlePheGlyHisCysProLeuTy 471
Db 1619 TGTTACACTGGCATCTGGCCGTGTGGTCATGTTGGTCATCTTCGGTCATTGCCCACTCTA 1678
Qy 471 rGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHi 491
Db 1679 TGGATATATAAGCGTTGTGCAGGAATATGACTTGGAAAAGAACACATGGAGTATATTACA 1738
Qy 491 sThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisSerSerValTyrAspHisArgTh 511
Db 1739 TACTCAGGGTGCTCTTGTGCAAGGGGGTTATGGCCACAGTAGTGTTTATGATGACAGGAC 1798
Qy 511 rArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAl 531
Db 1799 CAAGGCTCTGTACGTTCATGGTGGCTACAAGGCTTTCAGCGCCAACAAATACCGGCTTGC 1858
Qy 531 aAspAspLeuTyrArgTyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerAr 551
Db 1859 AGATGACCTCTACAGATACGATGTGGATACTCAGATGTGGACCATTCTTAAGGACAGCCG 1918
Qy 551 gPhePheArgTyrLeuHisThrAlaValIleValSerGlyThrMetLeuValPheGlyGl 571
Db 1919 ATTTTTCCGTTACTTGCATACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGAGG 1978
Qy 571 yAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMe 591
Db 1979 GAACACACACAATGACACTTCCATGAGCCACGGTGCCAAATGCTTCTCCTCGGACTTCAT 2038
Qy 591 tAlaTyrAspIleAlaCysAspArgTrpSerValLeuProArgProAspLeuHisHisAs 611
Db 2039 GGCTTATGACATTGCTTGTGACCGATGGTCAGTGCTTCCCAGACCTGAGCTCCATCATGA 2098
Qy 611 pValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGl 631
Db 2099 TGTCAACAGATTTGGCCATTCAGCAGTCTTGTACAACAGCACCATGTATGTGTTCGGCGG 2158
Qy 631 yPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHi 651
Db 2159 CTTCAACAGCCTCCTCCTCAGTGACGTCTTGGTCTTTACCTCGGAGCAGTGCGATGCACA 2218
Qy 651 sArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgCysValTrpAsnThrGl 671
Db 2219 CCGCAGTGAAGCTGCTTGTGTGGCAGCAGGACCTGGTATCCGGTGTCTGTGGGACACACA 2278
Qy 671 ySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnGluGluLysLeuLysSe 691
Db 2279 GTCGTCTCGATGTACCTCCTGGGAGTTGGCAACTGAAGAACAAGCAGAAAAGTTAAAATC 2338
Qy 691 rGluCysPheSerLysArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTy 711
Db 2339 AGAGTGTTTTTCTAAAAGAACCCTTGACCATGACAGATGTGACCAGCACACAGATTGTTA 2398
Qy 711 rSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAspHisCysValProArgAs 731
Db 2399 CAGCTGCACAGCCAATACCAATGACTGCCACTGGTGCAATGATCACTGTGTCCCTGTGAA 2458
Qy 731 nHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGluAsnCysProLysAspAs 751
Db 2459 CCACAGCTGCACAGAAGGCCAGATCTCCATTGCCAAGTATGAGAGTTGCCCCAAGGATAA 2518
Qy 751 nProMetTyrTyrCysAsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCy 771
Db 2519 CCCCATGTACTACTGCAATAAGAAAACCAGCTGCAGGAGCTGTGCCCTAGACCAGAACTG 2578
Qy 771 sGlnTrpGluProArgAsnGlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGl 791
Db 2579 CCAGTGGGAGCCCCGGAATCAAGAGTGCATCGCCCTGCCGGAAAATATCTGTGGCAATGG 2638
Qy 791 yTrpHisLeuValGlyAsnSerCysLeuLysIleThrThrAlaLysGluAsnTyrAspAs 811
Db 2639 CTGGCATTTGGTTGGAAACTCGTGTCTGAAAATCACTACTGCTAAGGAGAATTATGACAA 2698
Qy 811 nAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSerLeuThrThrGlnLysLy 831
Db 2699 TGCTAAATTGTCCTGTAGGAACCACAATGCCTTTTTGGCTTCCCTCACATCCCAGAAGAA 2758
Qy 831 sValGluPheValLeuLysGlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLe 851
Db 2759 GGTGGAGTTTGTCCTTAAGCAGCTTCGATTAATGCAATCATCTCAAAGTATGTCCAAGCT 2818
Qy 851 uThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAs 871
Db 2819 CACTCTGACTCCATGGGTTGGTCTTCGGAAGATCAATGTGTCTTACTGGTGCTGGGAGGA 2878
Qy 871 pMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSerAspAlaGl 891
Db 2879 TATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGTGATGCTGG 2938
Qy 891 yPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThrCysIleAs 911
Db 2939 CTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGATTAAAGGCTGCAACCTGCATCAA 2998
Qy 911 nProLeuAsnGlySerValCysGluArgProAlaAsnHisSerAlaLysGlnCysArgTh 931
Db 2999 CCCTCTCAATGGCAGCGTCTGTGAAAGGCCTGCAAACCACAGTGCCAAGCAGTGCCGGAC 3058
Qy 931 rProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTr 951
Db 3059 ACCATGTGCCCTGCGGACAGCGTGTGGCGAGTGCACTAGCAGCAGCTCGGAGTGCATGTG 3118
Qy 951 pCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrValAlaSerProProPheGl 971
Db 3119 GTGCAGTAACATGAAGCAGTGTGTGGACTCCAATGCCTACGTGGCCTCCTTCCCTTTTGG 3178
Qy 971 yGlnCysMetGluTrpTyrThrMetSerThrCysProProGluAsnCysSerGlyTyrCy 991
Db 3179 CCAGTGTATGGAATGGTATACGATGAGCAGCTGCCCACCTGAAAATTGCTCTGGCTACTG 3238
Qy 991 sThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCysThrAspProSerAsnTh 1011
Db 3239 TACCTGCAGCCATTGCTTGGAGCAGCCAGGCTGTGGTTGGTGTACTGATCCTAGCAATAC 3298
Qy 1011 rGlyLysGlyLysCysIleGluGlySerTyrLysGlyProValLysMetProSerGlnAl 1031
Db 3299 TGGGAAAGGAAAATGTATTGAGGGCAGCTATAAAGGACCTGTGAAGATGCCGTCACAGGC 3358
Qy 1031 aProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSe 1051
Db 3359 CTCTGCAGGAAATGTGTATCCACAGCCCCTTCTGAACTCCAGCATGTGTCTAGAGGACAG 3418
Qy 1051 rArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCy 1071
Db 3419 CAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAGTGCAACGGACACAGCAAATG 3478
Qy 1071 sIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrThrGlyLysHisCysGluTh 1091
Db 3479 CATCAACCAGAGTATCTGTGAGAAGTGTGAGGACCTGACCACGGGCAAGCACTGCGAGAC 3538
Qy 1091 rCysIleSerGlyPheTyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLysCy 1111
Db 3539 CTGCATATCTGGCTTCTATGGTGACCCGACTAATGGAGGCAAATGTCAGCCATGCAAGTG 3598
Qy 1111 sAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGl 1131
Db 3599 CAATGGGCACGCATCACTGTGCAACACCAACACCGGCAAGTGCTTCTGTACCACCAAAGG 3658
Qy 1131 yValLysGlyAspGluCysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLe 1151
Db 3659 TGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCT 3718
Qy 1151 uArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGl 1171
Db 3719 CAAAGGAACATGCTACTATACCCTTCTCATTGACTATCAGTTCACCTTTAGCCTGTCCCA 3778

```
Qy       1171 nGluAspAspArgTyrTyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnAr 1191
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3779 GGAAGACGACCGCTACTACACAGCCATCAACTTTGTGGCTACTCCTGATGAACAAAACAG 3838

Qy       1191 gAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAl 1211
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3839 GGATTTGGACATGTTCATCAATGCCTCCAAAAACTTCAACCTCAACATCACCTGGGCCAC 3898

Qy       1211 aSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIl 1231
             ||||||    ||||||||||||    |||||||||:::||||||||||||||||||||||||
Db       3899 CAGCTTCCCAGCCGGAACCCAGACTGGAGAAGAGGTGCCTGTTGTTTCAAAAACCAACAT 3958

Qy       1231 eLysGluTyrLysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIl 1251
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3959 CAAGGAATACAAAGATAGCTTCTCTAATGAGAAATTTGATTTTCGCAACCATCCAAACAT 4018

Qy       1251 eThrPhePheValTyrValSerAsnPheThrTrpProIleLysIleGlnIleAlaPheSe 1271
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4019 CACTTTCTTTGTTTATGTCAGTAATTTCACTTGGCCCATCAAAATTCAGATTGCCTTCTC 4078

Qy       1271 rGlnHisSerAsnPheMetAspLeuValGlnPhePheValThrPhePheSerCysPheLe 1291
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4079 CCAGCACAGCAACTTCATGGACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCT 4138

Qy       1291 uSerLeuLeuLeuValAlaAlaValValTrpLysIleLysGlnSerCysTrpAlaSerAr 1311
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4139 CTCGCTGCTTCTGGTGGCTGCAGTGGTCTGGAAGATCAAGCAGAGCTGTTGGGCATCCAG 4198

Qy       1311 gArgArgGluGlnLeuLeuArgGluMetGlnGlnMetAlaSerArgProPheAlaSerVa 1331
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4199 GCGGAGAGAGCAACTTCTTCGGGAGATGCAACAGATGGCCAGCCGCCCCTTTGCTTCTGT 4258

Qy       1331 lAsnValAlaLeuGluThrAspGluGluProProAspLeuIleGlyGlySerIleLysTh 1351
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4259 AAACGTTGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGAC 4318

Qy       1351 rValProLysProIleAlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerVa 1371
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4319 TGTTCCCAAACCCATTGCACTGGAGCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGT 4378

Qy       1371 lPheValArgLeuProArgGlyLeuGlyGlyIleProProProGlyGlnSerGlyLeuAl 1391
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4379 GTTTGTGAGGCTCCCTCGAGGCCTGGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGC 4438

Qy       1391 aValAlaSerAlaLeuValAspIleSerGlnGlnMetProIleValTyrLysGluLysSe 1411
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4439 TGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTC 4498

Qy       1411 rGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnProGlyThrCysIle 1429
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4499 AGGAGCCGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGCCTGGGACCTGCATC 4553

RESULT 3
US-09-893-238-16
; Sequence 16, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT:  Nagle, D.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
;  FILE REFERENCE: 7853-237
;  CURRENT APPLICATION NUMBER: US/09/893,238
;  CURRENT FILING DATE:  2001-06-27
;  PRIOR APPLICATION NUMBER: 09/245,041
;  PRIOR FILING DATE: 1999-02-05
;  PRIOR APPLICATION NUMBER: 60/093,630
;  PRIOR FILING DATE: 1998-07-21
;  PRIOR APPLICATION NUMBER: 60/104,978
;  PRIOR FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 16
;  LENGTH: 4072
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US-09-893-238-16

Alignment Scores:
Pred. No.:                0          Length:          4072
Score:                    6870.50    Matches:         1240
Percent Similarity:       97.41%     Conservative:    3
Best Local Similarity:    97.18%     Mismatches:      20
Query Match:              86.80%     Indels:          13
DB:                       9          Gaps:            3

US-09-787-097-12 (1-1429) x US-09-893-238-16 (1-4072)

Qy          2 ValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAlaAla 21
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         10 GTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCG 69

Qy         22 LeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArgPro 41
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 CTCGCGGGCAGGAGCGGCGGGCCGCACCGACCCTGCACCGCGACAGGGGCCTGGAGGCCG 129

Qy         42 GlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSer---------------ProPro 56
              |||   |||  |||  |||||||:::||||||                    ||||||
Db        130 GGACCGCGCGCCCGGCTGTGTCTCCCGCGGGTGCTGTCGCGGGCGCTGCCCCCGCCGCCG 189

Qy         57 LeuArgProArgLeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeu 76
              |||   |||   ||||||       ||||||||||||||||
Db        190 CTGCTGCCG---CTGCTCTTTTCGCTGCTGCTGCTGCCG---------------------- 225

Qy         77 LeuProCysGluAlaGluAlaAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAla 96
              ||||||   ||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db        226 CTGCCCCGGGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCC 285

Qy         97 GluAlaLysGluCysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGly 116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 GAGGCCAAGGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGC 345

Qy        117 GlnCysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPhe 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 CAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTC 405

Qy        137 ArgLeuThrGlySerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThr 156
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        406 AGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACG 465

Qy        157 LysCysThrTrpLeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHis 176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        466 AAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCAT 525

Qy        177 PheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAla 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        526 TTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCA 585

Qy        197 ProLeuValAlaAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrVal 216
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        586 CCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTC 645

Qy        217 ProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyr 236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        646 CCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTAT 705

Qy        237 AsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGly 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        706 AATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGC 765

Qy        257 ArgGlyGluCysLysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsn 276
              ||||||||||||||||||||||||||||:::||||||||||||||||||||||||||||||
Db        766 CGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAAC 825

Qy        277 TrpLysGlyGluAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArg 296
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        826 TGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGA 885
Qy        297 GlyIleCysAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyPro 316
Db        886 GGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCT 945
Qy        317 GlyCysSerValProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsn 336
Db        946 GGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAAC 1005
Qy        337 LeuLysLeuProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValVal 356
Db       1006 TTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTT 1065
Qy        357 GlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSer 376
Db       1066 GGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCT 1125
Qy        377 ArgGluTrpLeuProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSer 396
Db       1126 AGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCT 1185
Qy        397 LeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsn 416
Db       1186 TTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAAT 1245
Qy        417 ValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrPro 436
Db       1246 GTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCT 1305
Qy        437 LysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsn 456
Db       1306 AAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAAT 1365
Qy        457 GlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsn 476
Db       1366 GGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAAT 1425
Qy        477 ValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeu 496
Db       1426 GTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTT 1485
Qy        497 ValGlnGlyGlyTyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrVal 516
Db       1486 GTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTT 1545
Qy        517 HisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArg 536
Db       1546 CATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGA 1605
Qy        537 TyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeu 556
Db       1606 TATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTG 1665
Qy        557 HisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAsp 576
Db       1666 CACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGAC 1725
Qy        577 ThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAla 596
Db       1726 ACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCC 1785
Qy        597 CysAspArgTrpSerValLeuProArgProAspLeuHisHisAspValAsnArgPheGly 616
Db       1786 TGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGC 1845
Qy        617 HisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeu 636
Db       1846 CATTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTC 1905
Qy        637 LeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAla 656
Db       1906 CTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCT 1965

Qy        657 CysLeuAlaAlaGlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIle 676
Db       1966 TGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATC 2025
Qy        677 SerTrpAlaLeuAlaThrAspGluGlnGluGluLysLeuLysSerGluCysPheSerLys 696
Db       2026 TCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAA 2085
Qy        697 ArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsn 716
Db       2086 AGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAAC 2145
Qy        717 ThrAsnAspCysHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGlu 736
Db       2146 ACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAA 2205
Qy        737 GlyGlnIleSerIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCys 756
Db       2206 GGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGT 2265
Qy        757 AsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArg 776
Db       2266 AACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGG 2325
Qy        777 AsnGlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGly 796
Db       2326 AATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGA 2385
Qy        797 AsnSerCysLeuLysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCys 816
Db       2386 AACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGT 2445
Qy        817 ArgAsnHisAsnAlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeu 836
Db       2446 AGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTT 2505
Qy        837 LysGlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrp 856
Db       2506 AAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGG 2565
Qy        857 ValGlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThr 876
Db       2566 GTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACA 2625
Qy        877 AsnSerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeu 896
Db       2626 AATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTA 2685
Qy        897 SerGluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySer 916
Db       2686 TCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGT 2745
Qy        917 ValCysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArg 936
Db       2746 GTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGG 2805
Qy        937 ThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLys 956
Db       2806 ACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAG 2865
Qy        957 GlnCysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrp 976
Db       2866 CAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGG 2925
Qy        977 TyrThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCys 996
Db       2926 TATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGC 2985
Qy        997 LeuGluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCys 1016
Db       2986 TTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGC 3045
```

```
Qy       1017 IleGluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPhe 1036
Db       3046 ATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTC 3105

Qy       1037 TyrProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSer 1056
Db       3106 TATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCT 3165

Qy       1057 PheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIle 1076
Db       3166 TTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATC 3225

Qy       1077 CysGluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPhe 1096
Db       3226 TGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTC 3285

Qy       1097 TyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSer 1116
Db       3286 TACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCT 3345

Qy       1117 LeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGlu 1136
Db       3346 CTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAG 3405

Qy       1137 CysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyr 1156
Db       3406 TGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTAT 3465

Qy       1157 TyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyr 1176
Db       3466 TATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTAT 3525

Qy       1177 TyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPhe 1196
Db       3526 TACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTC 3585

Qy       1197 IleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGly 1216
Db       3586 ATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGA 3645

Qy       1217 ThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAsp 1236
Db       3646 ACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGAT 3705

Qy       1237 SerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyr 1256
Db       3706 AGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTAT 3765

Qy       1257 ValSerAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGln 1272
Db       3766 GTCAGTAATTTCACCTGGCCCATCAAAATTCAGGTGCAAACTGAACAA 3813
```

RESULT 4
US-10-197-824-1
; Sequence 1, Application US/10197824
; Publication No. US20040023219A1
; GENERAL INFORMATION:
; APPLICANT: ORIGENE TECHNOLOGIES INC
; TITLE OF INVENTION: NOVEL PROSTATE CANCER GENES
; FILE REFERENCE: 9U 106 R1
; CURRENT APPLICATION NUMBER: US/10/197,824
; CURRENT FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 6733
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (575)..(4711)
; OTHER INFORMATION:
US-10-197-824-1

Alignment Scores:

Pred. No.:	0	Length:	6733
Score:	4637.00	Matches:	853
Percent Similarity:	72.43%	Conservative:	211
Best Local Similarity:	58.07%	Mismatches:	310
Query Match:	58.58%	Indels:	95
DB:	16	Gaps:	26

US-09-787-097-12 (1-1429) x US-10-197-824-1 (1-6733)

```
Qy          6 AlaAlaThrGluAlaArgLeuArgArgArgThrAla----------Ala 18
Db        457 GCCGGTGAGGAGGAGGAGAAGCGGCGGCGGAGAGGTTTTCTGCGGCCGGAATTCCCTTCA 516

Qy         19 ThrAlaAlaLeu----------AlaGlyArgSer 26
Db        517 ACAGCATCCCTGTCGGCGCCCGCGAGCGCAGTCTCGCCGGGCAGGGGCGCCGGGGAAGAT 576

Qy         27 Gly----------GlyProHisTrpAspTrpAspValThrArgAlaGlyArgProGly 42
Db        577 GGAGACTGGGGGCCGGGCCCGCACTGG---------TACCCCGCAGCCAGCGGCCCCGGG 627

Qy         43 --------------LeuGlyAlaGlyLeuArgLeuProArgLeuLeuSerProProLeu 57
Db        628 GGTGTGGAGGGCTCGGCCGGCGGGCGGCGGCGGGGGGCGCCTCCTCCTGGCTGCTGGA 687

Qy         58 ArgProArgLeuLeuLeuLeuLeuLeuLeuProProProLeu-LeuLeuLeuLeLe 77
Db        688 CGGGAACAGCTGGCTGCTGTGCTATGGCTT---------CCTCTACCTGGCGCTCTAC-- 736

Qy         77 uProCysGluAlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAlaGl 97
Db        737 ----------------------------------------------GCGCAGGTGTCCCA 750

Qy         97 uAlaLysGluCysAspArg------ProCysValAsnGlyGlyArgCysAsnProGlyTh 115
Db        751 GTCCAAGCCGTGCGAGAGGACCGGCTCCTGCTTCTCG---GGCCGCTGTGTCAACTCCAC 807

Qy        115 rGlyGlnCysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyAr 135
Db        808 C------TGCCTCTGCGACCCGGGCTGGGTGGGGGACCAGTGCCAGCACTGCCAGGGCAG 861

Qy        135 gPheArgLeuThrGlySerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLy 155
Db        862 GTTCAAGTTAACAGAACCTTCTGGATATTTAACAGATGGCCCAATTAACTATAAATATAA 921

Qy        155 sThrLysCysThrTrpLeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAs 175
Db        922 AACTAAATGTACTTGGCTCATTGAAGGCTATCCAAATGCAGTGTTAAGATTAAGATTCAA 981

Qy        175 nHisPheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTy 195
Db        982 TCATTTTGCTACAGAATGTAGCTGGGATCATATGTATGTTTATGATGGAGATTCAATATA 1041

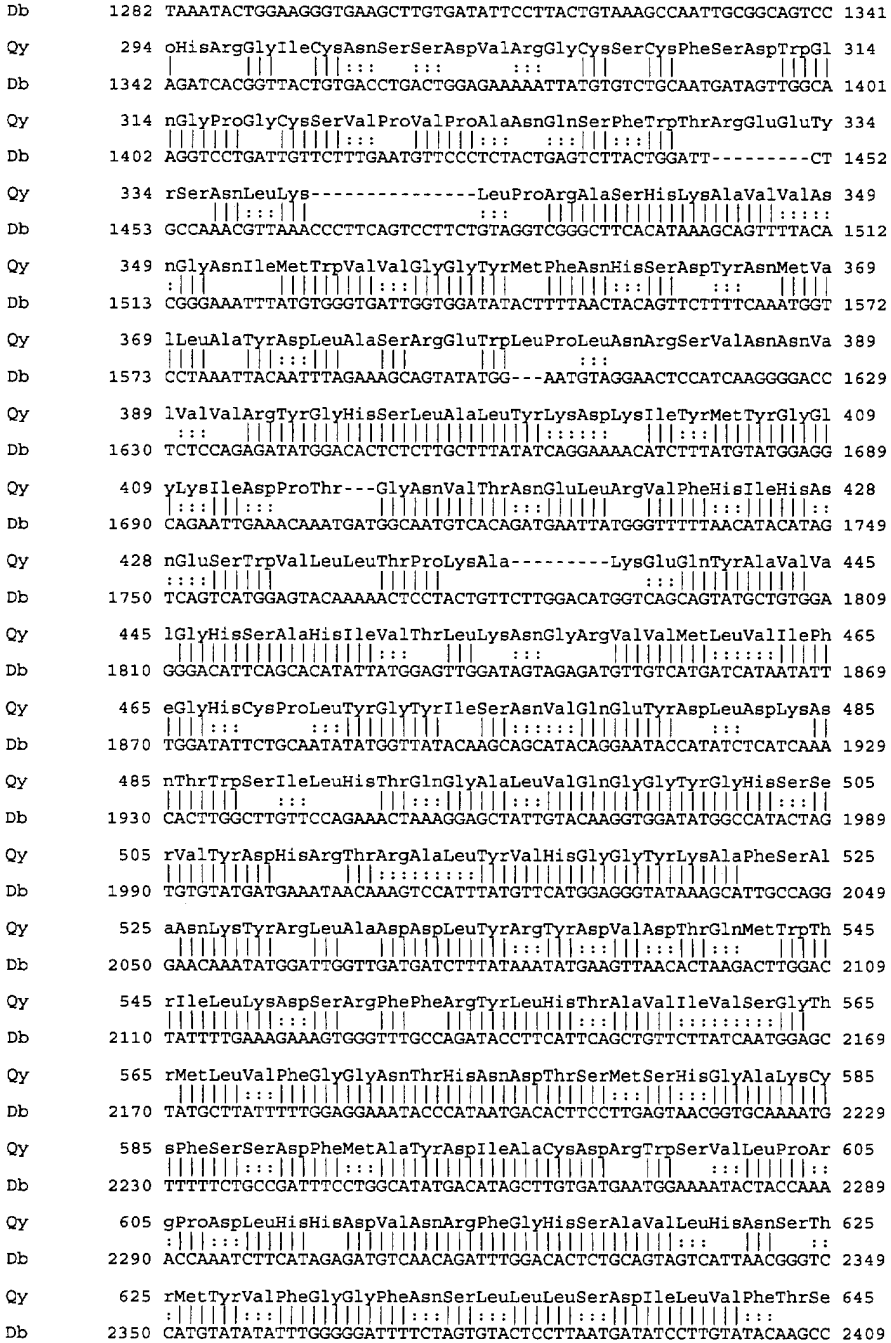
Qy        195 rAlaProLeuValAlaAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluTh 215
Db       1042 TGCACCTTTAATAGCTGTACTTAGTGGTTTGATAGTCCCTGAAATAAGGGGCAATGAAAC 1101

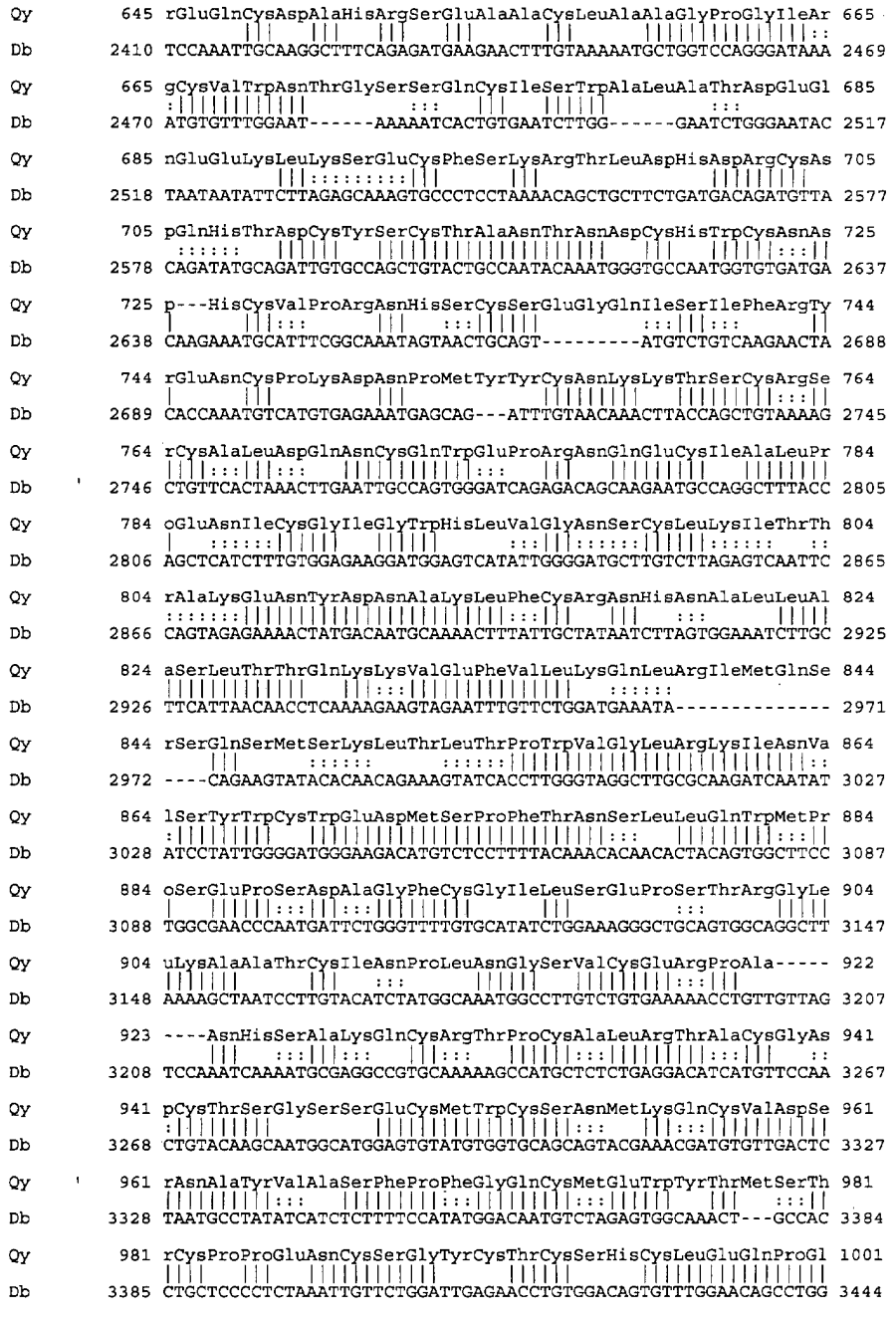
Qy        215 rValProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAl 235
Db       1102 TGTGCCTGAAGTTGTTACTACATCTGGCTATGCACTGTTACATTTTTTTAGTGATGCTGC 1161

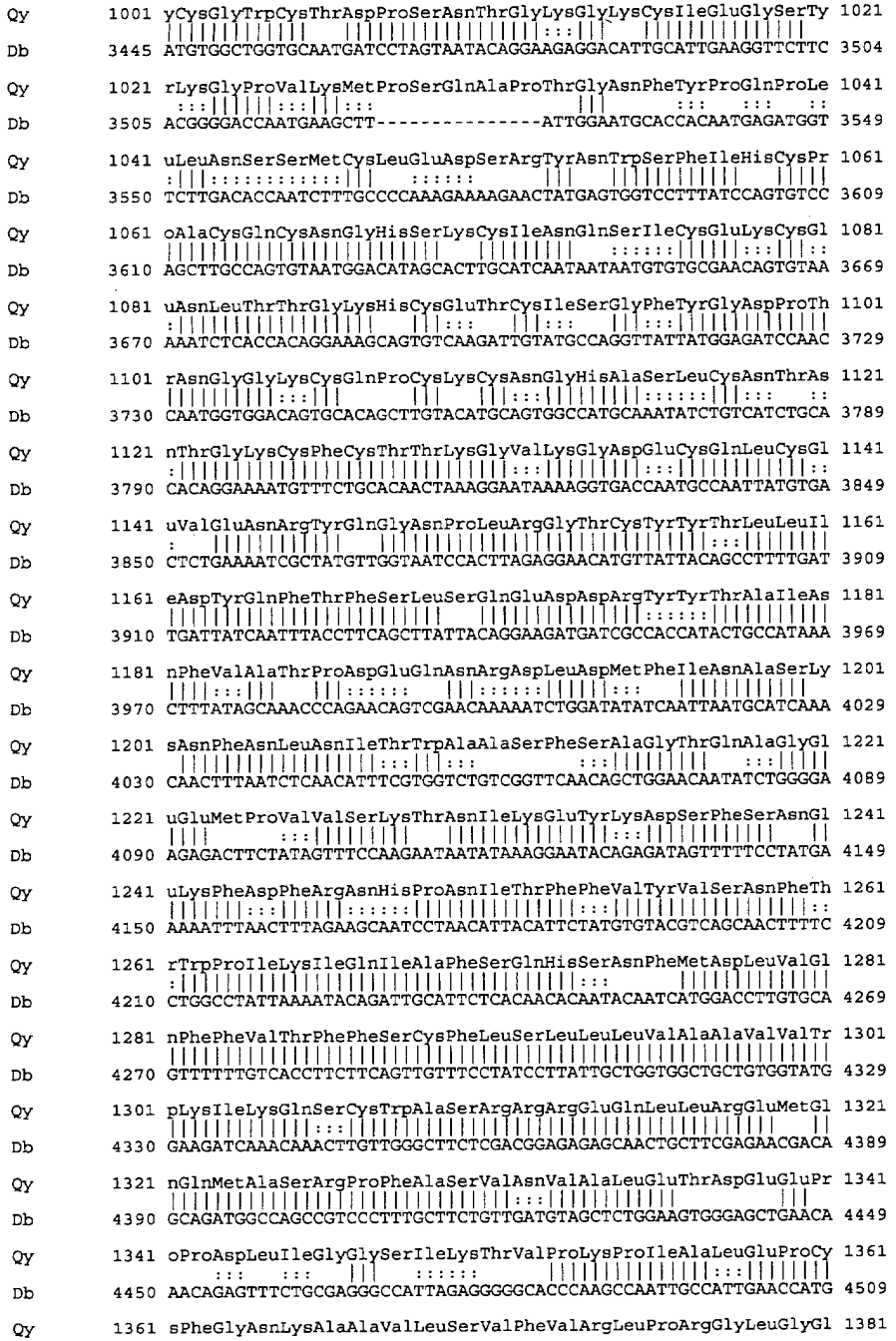
Qy        235 aTyrAsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSe 255
Db       1162 GTATAATCTAACTGGTTTCAACATTTTCTATTCAATCAATTCTTGTCCTAACAATTGCTC 1221

Qy        255 rGlyArgGlyGluCysLysIleSerAsnSer---SerGluThrValGluCysGluCysSe 274
Db       1222 TGGTCATGGGAAGTGTACAACTAGTGTCTCTGTTCCAAGTCAAGTATATTGTGAATGTGA 1281

Qy        274 rGluAsnTrpLysGlyGluAlaCysAspIleProHisCysThrAspAsnCysGlyPhePr 294
```





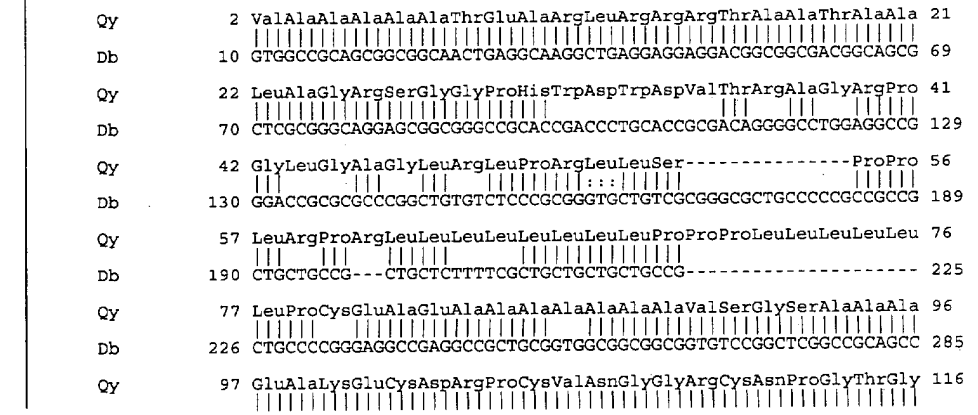


RESULT 5
US-09-893-238-18
; Sequence 18, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 2625
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-893-238-18

Alignment Scores:

Pred. No.:	0	Length:	2625
Score:	4136.50	Matches:	759
Percent Similarity:	95.48%	Conservative:	2
Best Local Similarity:	95.23%	Mismatches:	22
Query Match:	52.26%	Indels:	14
DB:	9	Gaps:	3

US-09-787-097-12 (1-1429) x US-09-893-238-18 (1-2625)



```
Db        286 GAGGCCAAGGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGC 345

Qy        117 GlnCysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPhe 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 CAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTC 405

Qy        137 ArgLeuThrGlySerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThr 156
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        406 AGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACG 465

Qy        157 LysCysThrTrpLeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHis 176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        466 AAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCAT 525

Qy        177 PheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAla 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        526 TTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCA 585

Qy        197 ProLeuValAlaAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrVal 216
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        586 CCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTC 645

Qy        217 ProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyr 236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        646 CCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTAT 705

Qy        237 AsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGly 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        706 AATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGC 765

Qy        257 ArgGlyGluCysLysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsn 276
              |||||||||||||||||||||||||||||||:::||||||||||||||||||||||||||
Db        766 CGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAAC 825

Qy        277 TrpLysGlyGluAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArg 296
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        826 TGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGA 885

Qy        297 GlyIleCysAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyPro 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        886 GGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCT 945

Qy        317 GlyCysSerValProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsn 336
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        946 GGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAAC 1005

Qy        337 LeuLysLeuProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValVal 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1006 TTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTT 1065

Qy        357 GlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSer 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1066 GGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCT 1125

Qy        377 ArgGluTrpLeuProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSer 396
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1126 AGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCT 1185

Qy        397 LeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsn 416
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1186 TTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAAT 1245

Qy        417 ValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrPro 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1246 GTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCT 1305

Qy        437 LysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsn 456
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1306 AAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAAT 1365

Qy        457 GlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsn 476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1366 GGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAAT 1425

Qy        477 ValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeu 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1426 GTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTT 1485

Qy        497 ValGlnGlyGlyTyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrVal 516
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1486 GTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTT 1545

Qy        517 HisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArg 536
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1546 CATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGA 1605

Qy        537 TyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeu 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1606 TATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTG 1665

Qy        557 HisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAsp 576
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1666 CACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGAC 1725

Qy        577 ThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAla 596
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1726 ACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCC 1785

Qy        597 CysAspArgTrpSerValLeuProArgProAspLeuHisHisAspValAsnArgPheGly 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1786 TGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGC 1845

Qy        617 HisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeu 636
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1846 CATTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTC 1905

Qy        637 LeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAla 656
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1906 CTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCT 1965

Qy        657 CysLeuAlaAlaGlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIle 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1966 TGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATC 2025

Qy        677 SerTrpAlaLeuAlaThrAspGluGlnGluGluLysLeuLysSerGluCysPheSerLys 696
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2026 TCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAA 2085

Qy        697 ArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsn 716
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2086 AGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAAC 2145

Qy        717 ThrAsnAspCysHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGlu 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2146 ACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAA 2205

Qy        737 GlyGlnIleSerIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCys 756
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2206 GGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGT 2265

Qy        757 AsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArg 776
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2266 AACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGG 2325

Qy        777 AsnGlnGluCysIleAlaLeuProGluAsnIleCys-GlyIleGlyTrp 792
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db       2326 AATCAGGAGTGCATTGCCCTGCCCGGTAGGCCTTGCAGGGTCATCTTGG 2374

RESULT 6
US-09-893-238-12
; Sequence 12, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT:  Nagle, D.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
```

```
;  TITLE OF INVENTION:  DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
;  FILE REFERENCE: 7853-237
;  CURRENT APPLICATION NUMBER: US/09/893,238
;  CURRENT FILING DATE:  2001-06-27
;  PRIOR APPLICATION NUMBER: 09/245,041
;  PRIOR FILING DATE: 1999-02-05
;  PRIOR APPLICATION NUMBER: 60/093,630
;  PRIOR FILING DATE: 1998-07-21
;  PRIOR APPLICATION NUMBER: 60/104,978
;  PRIOR FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 12
;   LENGTH: 6370
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-893-238-12

Alignment Scores:
Pred. No.:                0              Length:          6370
Score:                    3491.50        Matches:         656
Percent Similarity:       95.51%         Conservative:    3
Best Local Similarity:    95.07%         Mismatches:      18
Query Match:              44.11%         Indels:          13
DB:                       9              Gaps:            4

US-09-787-097-12 (1-1429) x US-09-893-238-12 (1-6370)

Qy        753 MetTyrTyrCysAsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGln 772
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAG 60

Qy        773 TrpGluProArgAsnGlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGlyTrp 792
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGG 120

Qy        793 HisLeuValGlyAsnSerCysLeuLysIleThrThrAlaLysGluAsnTyrAspAsnAla 812
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCT 180

Qy        813 LysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSerLeuThrThrGlnLysLysVal 832
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTA 240

Qy        833 GluPheValLeuLysGlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThr 852
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACC 300

Qy        853 LeuThrProTrpValGly---LeuArgLysIleAsnValSerTyr-TrpCysTrp-GluA 871
              |||||||||||||||||     :::   ||||||||||||    |||||| |||
Db        301 TTAACCCCATGGGTCGGGCCTTCGGGAAGGTYCAATGTGTCCTACTKGGTGCTGGGGAAG 360

Qy        871 spMetSerPro---PheThrAsnSer-LeuLeuGlnTrpMetProSerGluProSerAsp 889
                  |||    |||||||||||| ||||||||||||                  :::
Db        361 GATATGKTCCCCATTTTACAAATAGTTTTACTACAGTGGGATGSCCGTCTTGAGGCCCAG 420

Qy        890 Ala------GlyPheCysGly---IleLeuSerGluProSer--ThrArgGlyLeuLysA 906
                      |||||||||     :::   |||||||||    ||||||||||||||
Db        421 TGTTGCTTGGRATTCTGTGGGAATTTTATTCAGGAACCCAGTTACTTCGGGGACTGAAGG 480

Qy        906 laAlaThrCysIle-AsnProLeu-AsnGlySerValCysGluArgProAlaAsnHisSe 925
              |||||||||||||| ||||||||| ||||||||||||||||||||||||||||||||||||
Db        481 CTGCAACCTGCATTCAACCCACTYMAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAG 540

Qy        925 rAlaLys-GlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThrSerG 945
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGCTAAGGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCG 600

Qy        945 lySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrV 965
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATG 660

Qy        965 alAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSerThrCysProProG 985
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTG 720

Qy        985 luAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpC 1005
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGT 780

Qy       1005 ysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyProV 1025
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAG 840

Qy       1025 alLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerS 1045
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCA 900

Qy       1045 erMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnC 1065
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAAT 960

Qy       1065 ysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrT 1085
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCA 1020

Qy       1085 hrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyGlyL 1105
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGA 1080

Qy       1105 ysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysC 1125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGT 1140

Qy       1125 ysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGluAsnA 1145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATC 1200

Qy       1145 rgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrGlnP 1165
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGT 1260

Qy       1165 heThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIleAsnPheValAlaT 1185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTA 1320

Qy       1185 hrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnL 1205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACC 1380

Qy       1205 euAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProV 1225
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTG 1440

Qy       1225 alValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPheAspP 1245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATT 1500

Qy       1245 heArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpProIleL 1265
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCA 1560

Qy       1265 ysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePheValT 1285
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCTGGTACAGTTCTTCGTGA 1620

Qy       1285 hrPhePheSerCysPheLeuSerLeuLeuLeuValAlaAlaValValTrpLysIleLysG 1305
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGTGGTTTGGAAGATCAAAC 1680

Qy       1305 lnSerCysTrpAlaSerArgArgArgGluGlnLeuLeuArgGluMetGlnGlnMetAlaS 1325
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGAGATGCAACAGATGGCCA 1740

Qy       1325 erArgProPheAlaSerValAsnValAlaLeuGluThrAspGluGluProProAspLeuI 1345
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1741 GCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTA 1800

Qy       1345 leGlyGlySerIleLysThrValProLysProIleAlaLeuGluProCysPheGlyAsnL 1365
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 TTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGAGCCGTGTTTTGGCAACA 1860

Qy       1365 ysAlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyGlyIleProProP 1385
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 AAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCTGGGTGGCATCCCTCCTC 1920

Qy       1385 roGlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetProI 1405
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 CTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGA 1980

Qy       1405 leValTyrLysGluLysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnP 1425
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGC 2040

Qy       1425 roGlyThrCysIle 1429
              |||||||||||||
Db       2041 CTGGGACCTGCATC 2054

RESULT 7
US-09-893-238-8
; Sequence 8, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT:  Nagle, D.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
;  FILE REFERENCE: 7853-237
;  CURRENT APPLICATION NUMBER: US/09/893,238
;  CURRENT FILING DATE:  2001-06-27
;  PRIOR APPLICATION NUMBER: 09/245,041
;  PRIOR FILING DATE: 1999-02-05
;  PRIOR APPLICATION NUMBER: 60/093,630
;  PRIOR FILING DATE: 1998-07-21
;  PRIOR APPLICATION NUMBER: 60/104,978
;  PRIOR FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
;   LENGTH: 2419
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-893-238-8

Alignment Scores:
Pred. No.:                0              Length:          2419
Score:                    3435.00        Matches:         642
Percent Similarity:       84.37%         Conservative:    22
Best Local Similarity:    81.58%         Mismatches:      44
Query Match:              43.40%         Indels:          79
DB:                       9              Gaps:            4

US-09-787-097-12 (1-1429) x US-09-893-238-8 (1-2419)

Qy          2 ValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAlaAla 21
              ||||||||||||||||||||||||||||||||||||      |||      |||||||||
Db         83 GTGGCGGCGGCGGCGGCGACTGAGGCGCGGCTGAGGGGGAGCACGAGGACGACAGCAGCG 142

Qy         22 LeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArgPro 41
                 |||||||||   |||      |||                |||   |||   ||||||
Db        143 CCTGCGGGCAGGAAGGGCAGGCAGCACCGACCCTGCACCGCGACAGGGGCCTGGAGGCCG 202

Qy         42 GlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSer----------------ProPro 56
              |||      |||   |||   |||||||||:::||||||                ||||||
Db        203 GGACCGCGCGCCCGGCTGTGTCTCCCGCGGGTGCTGTCGCGGGCGCTGCCCCCGCCGCCG 262

Qy         57 LeuArgProArgLeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeu 76
              |||   |||   ||||||      ||||||||||||||
Db        263 CTGCTGCCG---CTGCTCTTTTCGCTGCTGCTGCTGCCG---------------------- 298

Qy         77 LeuProCysGluAlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAla 96
              ||||||   ||||||||||||||||||   ||||||||||||||||||||||||||||||
Db        299 CTGCCCCGGGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCC 358

Qy         97 GluAlaLysGluCysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGly 116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 GAGGCCAAGGAATGTGACCGGCCGTGTGTCAACGGCGGCCGCTGCAACCCTGGCACCGGC 418

Qy        117 GlnCysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPhe 136
              ||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        419 CAGTGCGTCTGCCCCACGGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTC 478

Qy        136 ------------------------------------------------------------ 136

Db        479 AGGACATCTGTCTCACGCCTATAATCACAGCTGTTCGGAAGGTGAGGCTGGAGGAACAGT 538

Qy        136 ------------------------------------------------------------ 136

Db        539 TCGAGGCAAGCTTCGGCTACAGAATAAGTTCAAGAGTAACCTGGGGCAACTTGGGCTTGT 598

Qy        136 ------------------------------------------------------------ 136

Db        599 CTCCAAAACCAAAATGAGCGAAAAGGAGCAAGCTAGAGTCTTTTGGGAAAATTTTAGCTG 658

Qy        137 ---------------Arg-LeuThrGlySerSerGlyPheValThrAspGlyProGlyAs 151
                             ||| ||||||||||||||||||||||||||||||||||||||||||
Db        659 ACTAATTTTTCACCGAGAACTAACTGGCTCTTCTGGATTTGTAACAGATGGACCTGGGAA 718

Qy        151 nTyrLysTyrLysThrLysCysThrTrpLeuIleGluGlyGlnProAsnArgIleMetAr 171
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 TTATAAATATAAGACGAAGTGCACATGGCTCATTGAAGGACAGCCAAATAGAATAATGAG 778

Qy        171 gLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGl 191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 ACTTCGCTTCAACCATTTTGCTACAGAATGTAGCTGGGACCATTTATATGTTTATGATGG 838

Qy        191 yAspSerIleTyrAlaProLeuValAlaAlaPheSerGlyLeuIleValProGluArgAs 211
              ||||||||||||||||||||||:::||||||||||||||||||||||||||||||||||||
Db        839 GGACTCAATCTACGCACCTCTGATTGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGA 898

Qy        211 pGlyAsnGluThrValProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePh 231
              |||||||||||||   |||||||||      ||||||||||||||||||||||||||||||
Db        899 TGGCAATGAGACGGCTCCTGAGGTCACTGTCACTTCAGGTTATGCACTGCTGCATTTTTT 958

Qy        231 eSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysPr 251
              ||||||||||||||||||||||||||||||||||||||||||:::||||||||||||||||
Db        959 CAGTGATGCTGCTTATAATCTGACTGGATTTAATATCACTTACAATTTTGACATGTGTCC 1018

Qy        251 oAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSerSerGluThrValGluCy 271
              ||||||||||||||||||||||||||||||   ||||||||||      |||||||
Db       1019 GAATAATTGCTCAGGCCGAGGAGAGTGTAAGAGCAGTAACAGCAGCAGCGCTGTTGAGTG 1078

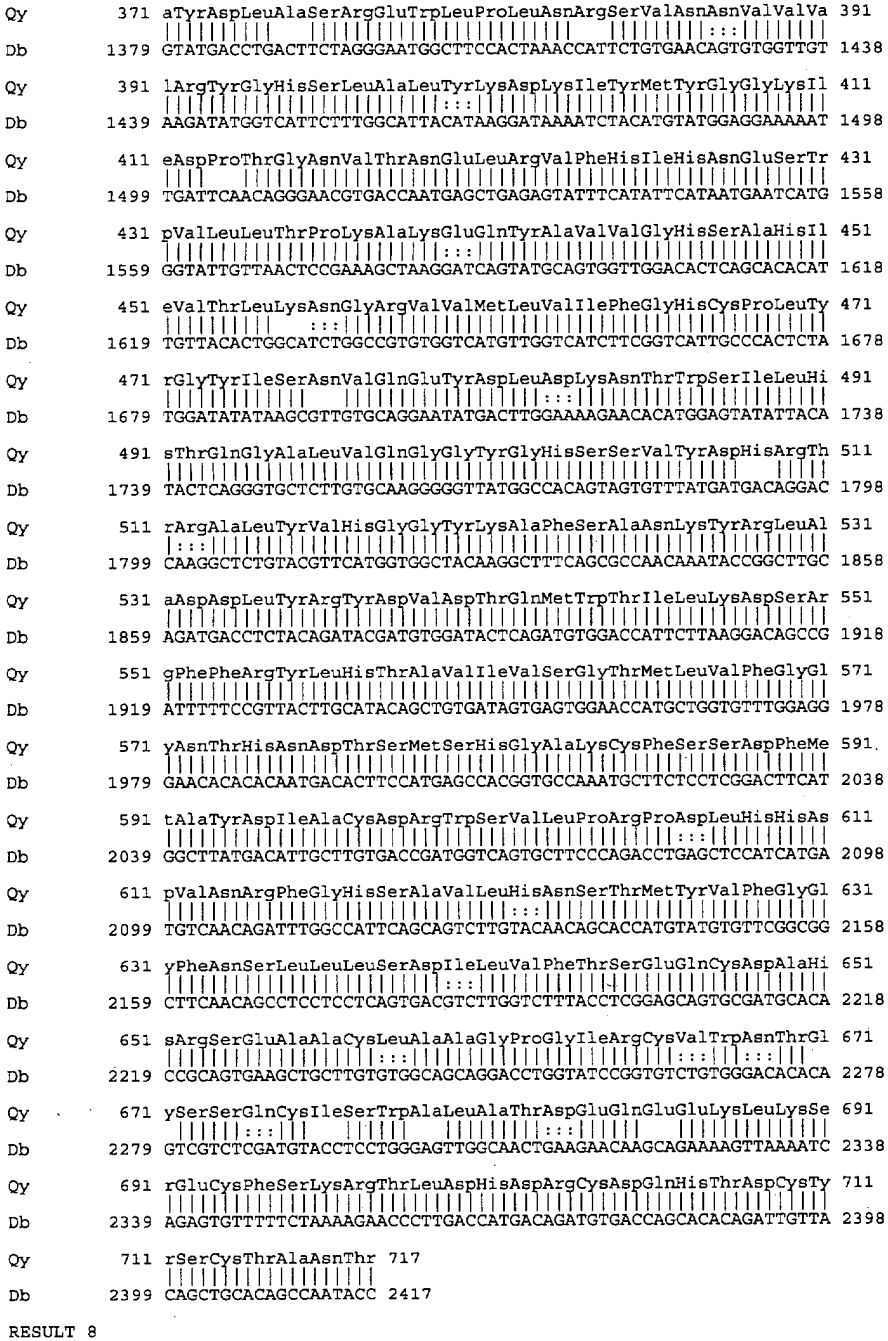
Qy        271 sGluCysSerGluAsnTrpLysGlyGluAlaCysAspIleProHisCysThrAspAsnCy 291
              ||||||||||||||||||||||||||:::||||||||||||||||||||||||||||||||
Db       1079 TGAATGTTCTGAAAACTGGAAAGGGGAGTCGTGTGACATTCCTCACTGTACAGACAACTG 1138

Qy        291 sGlyPheProHisArgGlyIleCysAsnSerSerAspValArgGlyCysSerCysPheSe 311
              ||||||||||||||||||||||||||:::||||||  ||||||||||||||||||
Db       1139 TGGCTTTCCTCACCGAGGCATCTGTAATGCAAGCGATACCAGAGGGTGCTCCTGCTTTCC 1198

Qy        311 rAspTrpGlnGlyProGlyCysSerValProValProAlaAsnGlnSerPheTrpThrAr 331
                 ||||||||||||||||||||:::||||||||||||||||||||||||||||||||||||
Db       1199 TCACTGGCAGGGTCCTGGATGTTCAATTCCTGTGCCAGCTAACCAGTCTTTTTGGACTCG 1258

Qy        331 gGluGluTyrSerAsnLeuLysLeuProArgAlaSerHisLysAlaValValAsnGlyAs 351
              ||||||||||||||:::||||||||||||||||||||||||||||||||||||||||||||
Db       1259 AGAAGAATATTCTGATTTAAAGCTTCCCAGAGCCTCTCATAAAGCTGTGGTCAATGGAAA 1318

Qy        351 nIleMetTrpValValGlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAl 371
              ||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||
Db       1319 TATAATGTGGGTTGTTGGCGGATATATGTTCAACCATTCAGATTACAGCATGGTTTTAGC 1378
```



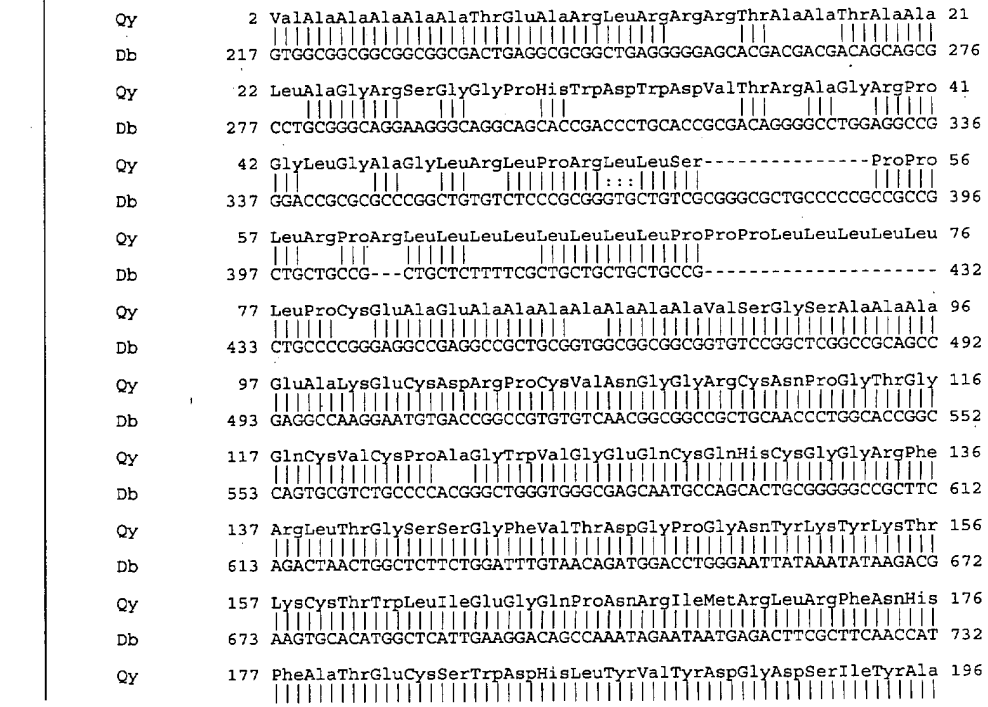
RESULT 8

US-09-893-238-10
; Sequence 10, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 1051
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-893-238-10

Alignment Scores:

Pred. No.:	2.18e-110	Length:	1051
Score:	1226.50	Matches:	236
Percent Similarity:	84.45%	Conservative:	3
Best Local Similarity:	83.39%	Mismatches:	31
Query Match:	15.50%	Indels:	13
DB:	9	Gaps:	3

US-09-787-097-12 (1-1429) x US-09-893-238-10 (1-1051)



```
Db        733 TTTGCTACAGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATCTACGCA 792

Qy        197 ProLeuValAlaAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrVal 216
              ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db        793 CCTCTGATTGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCT 852

Qy        217 ProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyr 236
              |||||||||     ||||||||||||||||||||||||||||||||||||||||||||||
Db        853 CCTGAGGTCACTGTCACTTCAGGTTATGCACTGCTGCATTTTTTCAGTGATGCTGCTTAT 912

Qy        237 AsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGly 256
              |||||||||||||||||||||||||:::|||||||||||||||||||||||||||||||
Db        913 AATCTGACTGGATTTAATATCACTTACAATTTTGACATGTGTCCGAATAATTGCTCAGGC 972

Qy        257 ArgGlyGluCysLysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsn 276
              |||||||||||||||  ||||||||||      ||||||||||||||||||||||||||
Db        973 CGAGGAGAGTGTAAGAGCAGTAACAGCAGCAGCGCTGTTGAGTGTGAATGTTCTGAAAAC 1032

Qy        277 TrpLysGly 279
              |||||||||
Db       1033 TGGAAAGGG 1041

RESULT 9
US-10-085-198-15
; Sequence 15, Application US/10085198
; Publication No. US20040009907A1
; GENERAL INFORMATION:
;  APPLICANT: Alsobrook et al.
;  TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 21402-279
;  CURRENT APPLICATION NUMBER: US/10/085,198
;  CURRENT FILING DATE:  2002-02-25
;  PRIOR APPLICATION NUMBER: 60/271,646
;  PRIOR FILING DATE: 2001-02-26
;  PRIOR APPLICATION NUMBER: 60/276,401
;  PRIOR FILING DATE: 2001-03-16
;  PRIOR APPLICATION NUMBER: 60/311,981
;  PRIOR FILING DATE: 2001-08-13
;  PRIOR APPLICATION NUMBER: 60/312,858
;  PRIOR FILING DATE: 2001-08-16
;  PRIOR APPLICATION NUMBER: 60/271,840
;  PRIOR FILING DATE: 2001-02-27
;  PRIOR APPLICATION NUMBER: 60/277,324
;  PRIOR FILING DATE: 2001-03-20
;  PRIOR APPLICATION NUMBER: 60/286,096
;  PRIOR FILING DATE: 2001-04-21
;  PRIOR APPLICATION NUMBER: 60/299,695
;  PRIOR FILING DATE: 2001-06-20
;  PRIOR APPLICATION NUMBER: 60/315,614
;  PRIOR FILING DATE: 2001-08-29
;  PRIOR APPLICATION NUMBER: 60/272,405
;  PRIOR FILING DATE: 2001-02-28
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 653
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
;   LENGTH: 8589
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-085-198-15

Alignment Scores:
Pred. No.:                9.96e-83      Length:          8589
Score:                    960.50        Matches:         333
Percent Similarity:       36.49%        Conservative:    152
Best Local Similarity:    25.06%        Mismatches:      496
Query Match:              12.14%        Indels:          350
DB:                       15            Gaps:            55

US-09-787-097-12 (1-1429) x US-10-085-198-15 (1-8589)

Qy        104 ProCysValAsnGlyGlyArgCysAsnProGly------ThrGlyGlnCysValCysPro 121
              ||||||         |||||||||   ||||||         |||   |||

Db         61 CCGTGC------TGGGGTCGCTGTCC-CCTGGGGCCCGGGCGGGGGACTGCAAG------ 107

Qy        122 AlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySer 141
                                                  |||   |||     |||     :::
Db        108 --------------------------------GGGCAGCGGCAGGTGCTGCGGGAGGCG 134

Qy        142 SerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeu 161
              |||||||||||||||||||  |||||||||                 |||   ||||||
Db        135 CCAGGCTTCGTGACGGATGGTGCGGGCAACTACAGCGTCAATGGCAACTGCGAGTGGCTC 194

Qy        162 IleGluGly---GlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGlu 180
              ||||||       |||     :::   |||   |||             ||||||
Db        195 ATCGAGGCCCCAAGCCCCCAGCACCGGATCCTGCTGGACTTCCTTTTCCTGGACACAGAG 254

Qy        181 CysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAla 200
              |||::::::|||:::|||:::|||||||||||||||||      ||||||:::|||
Db        255 TGCACGTATGACTACCTGTTCGTGTATGACGGTGACTCCCCGCGAGGGCCGCTGCTTGCC 314

Qy        201 AlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGluValVal 220
              :::   ||||||        |||                         |||   :::
Db        315 AGTCTAAGTGGGAGCACCCGACCT----------------------CCGCCCATCGAA 350

Qy        221 AlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGly 240
              |||:::||||||      |||||||||   ||||||||||   |||||||||   |||
Db        351 GCTTCCTCAGGCAAGATGCTGCTGCACCTCTTCAGTGATGCCAACTACAACCTGCTGGGC 410

Qy        241 PheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCys 260
              ||||||   ::::::   |||   :::|||||||        |||         |||:::|||
Db        411 TTTAACGCCTCATTCCGCTTCTCCCTGTGCCCGGGTGGCTGCCAGAGCCACGGGCAGTGC 470

Qy        261 LysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsnTrpLysGlyGlu 280
              :::                     |||  |||         |||   |||
Db        471 CAGCCACCGGGT--------------GTGTGTGCCTGCGAGCCGGGCTGGGGGGGTCCT 515

Qy        281 AlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsn 300
              |||   :::      |||:::     ||||||         |||   |||
Db        516 GACTGTGGCCTGCAGGAGTGCTCAGCCTACTGTGGC------AGCCACGGCACCTGC--- 566

Qy        301 SerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysServal 320
              :::|||   :::      |||   |||         :::   |||      |||   :::
Db        567 GCCTCGCCCCTGGGACCATGCCGCTGTGAGCCTGGCTTCTTGGGACGTGCCTGTGACCTG 626

Qy        321 ProValProAlaAsnGln------SerPheTrp-----------ThrArgGluGluTyr 334
                 :::      ||||||        :::|||                |||:::
Db        627 CACCTGTGGGAGAACCAGGGGGCTGGGTGGTGGCACAACGTGAGTGCCAGGGACCCTGCC 686

Qy        335 SerAsnLeuLysLeuProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrp 354
                 :::  ::::::      |||                |||   ::::::
Db        687 TTCTCTGCCCGTATTGGGGCAGCTGGCGCCTTCCTGTCCCCACCAGGG---CTGCTGGCA 743

Qy        355 ValValGlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeu 374
              |||   ||||||        |||::::::     ::::::||||||   |||:::
Db        744 GTTTTCGGAGGCCAGGACCTCAACAATGCCCTGGGTGACCTCGTCCTA---TACAACTTC 800

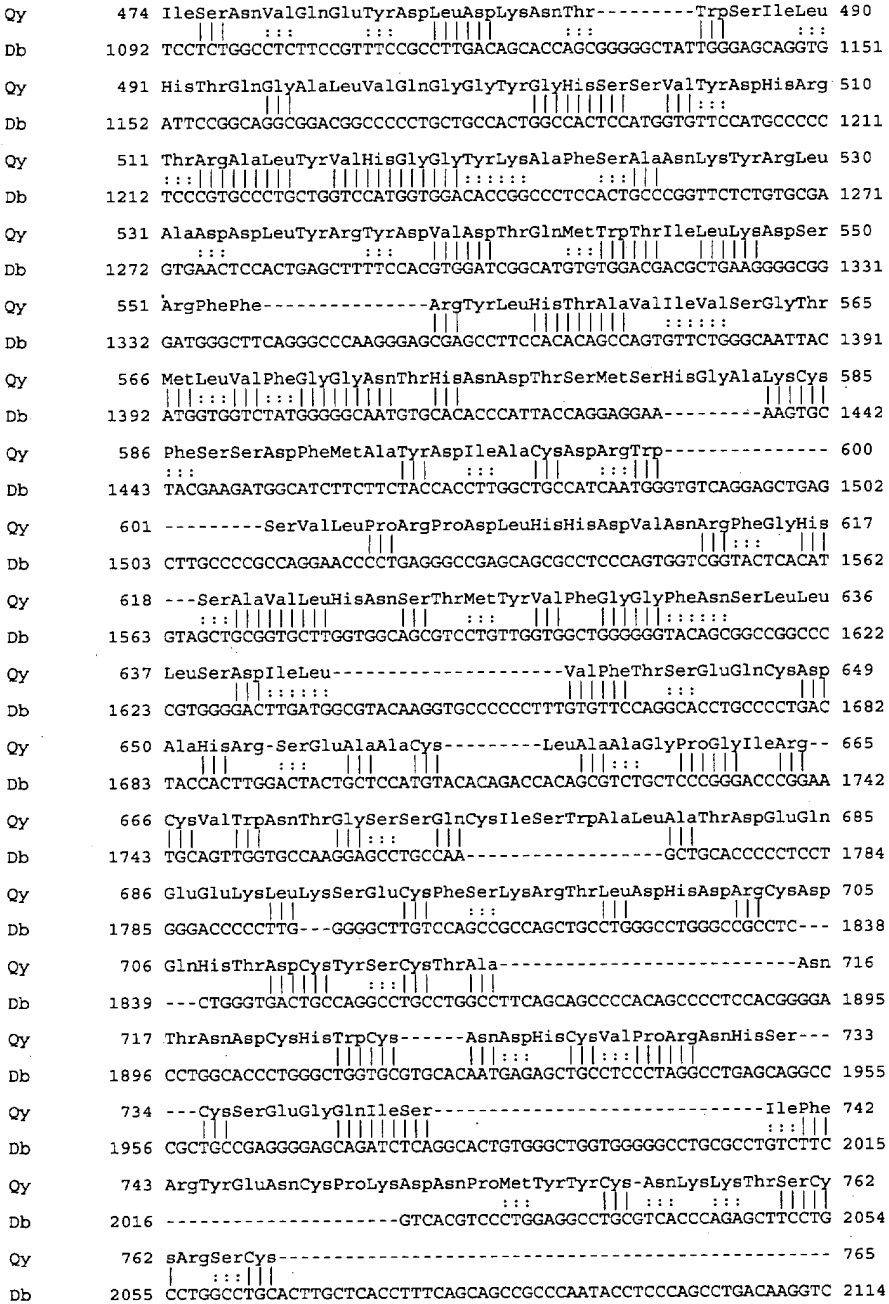
Qy        375 AlaSerArgGluTrpLeuProLeuAsnArgSerValAsnAsnValValValArgTyrGly 394
                 :::::      |||         :::   |||                |||:::
Db        801 TCCGCCAACACCTGGGAGTCTTGGGACCTGAGT---CCTGCCCCGGCTGCCCGTCACTCC 857

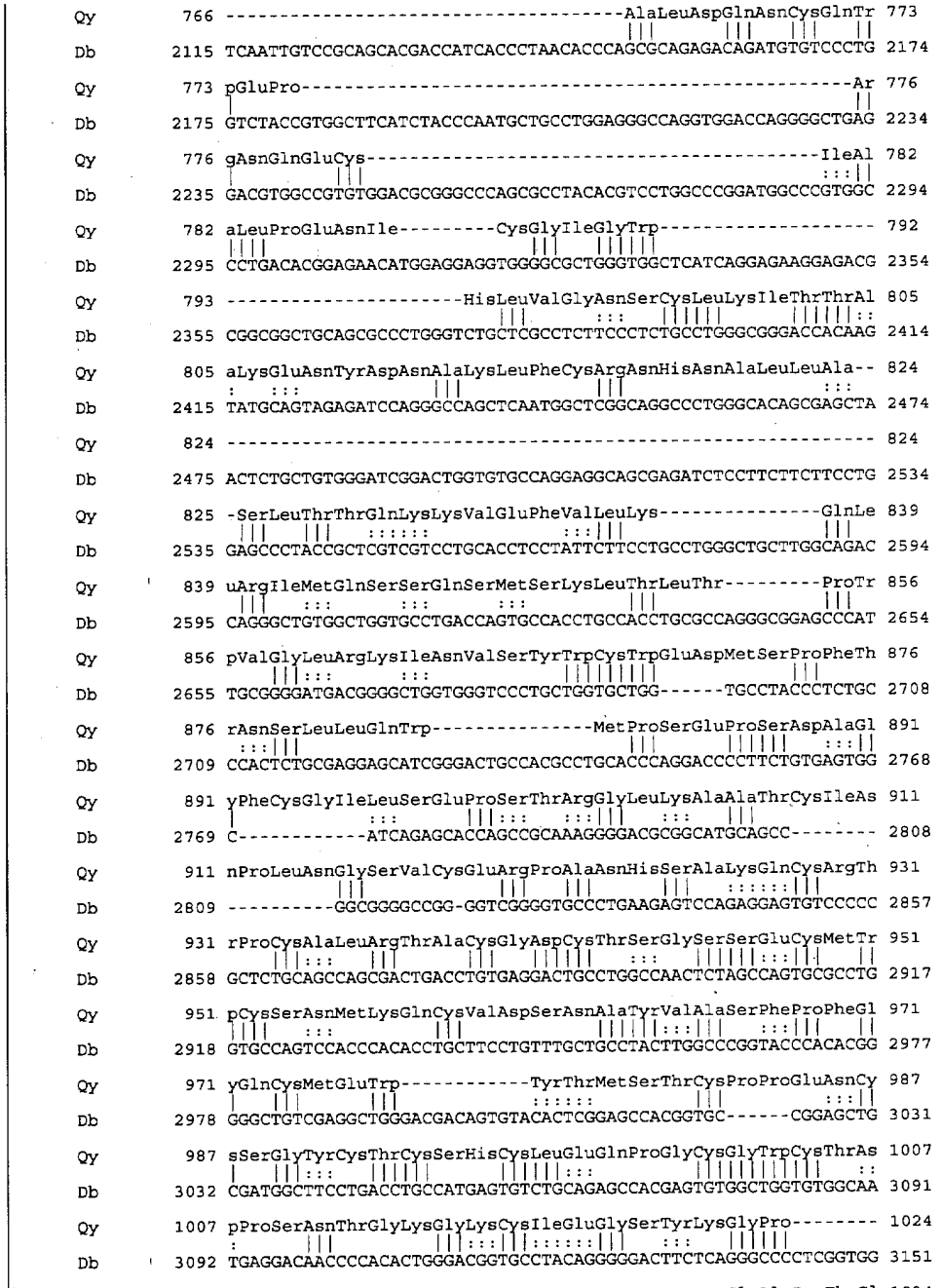
Qy        395 HisSerLeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLysIleAspProThr 414
              |||             :::       :::   :::   ||||||::::::
Db        858 CATGTGGCCGTGGCCTGGGCCGGCTCCCTGGTACTGATGGGTGGTGAGCTG---GCTGAC 914

Qy        415 GlyAsnValThrAsnGluLeuArgValPheHis---IleHisAsnGluSerTrpValLeu 433
              |||::::::|||||::::::     |||     :::           |||   |||
Db        915 GGCTCGCTCACCAACGACGTGTGGGCCTTCAGTCCACTGGGCAGGGGCCACTGGGAGCTC 974

Qy        434 LeuThrProLysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThr 453
              |||   |||   |||                     |||   :::   |||
Db        975 CTGGCACCACCTGCCTCCAGCTCCTCGGGGCCCCCAGGCCTGGCAGGTCACGCGGCTGCC 1034

Qy        454 LeuLysAsnGlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyr 473
              |||   :::      |||   :::   |||      |||      |||         :::
Db       1035 CTGGTGGATGAT---GTCTGGCTATATGTGTCTGGAGGCCGCACCCCGCACGACCTCTTC 1091
```


Qy 474 IleSerAsnValGlnGluTyrAspLeuAspLysAsnThr---------TrpSerIleLeu 490
Db 1092 TCCTCTGGCCTCTTCCGTTTCCGCCTTGACAGCACCAGCGGGGGCTATTGGGAGCAGGTG 1151
Qy 491 HisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisSerSerValTyrAspHisArg 510
Db 1152 ATTCCGGCAGGCGGACGGCCCCCTGCTGCCACTGGCCACTCCATGGTGTTCCATGCCCCC 1211
Qy 511 ThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeu 530
Db 1212 TCCCGTGCCCTGCTGGTCCATGGTGGACACCGGCCCTCCACTGCCCGGTTCTCTGTGCGA 1271
Qy 531 AlaAspAspLeuTyrArgTyrAspValAspThrGlnMetTrpThrIleLeuLysAspSer 550
Db 1272 GTGAACTCCACTGAGCTTTTCCACGTGGATCGGCATGTGTGGACGACGCTGAAGGGGCGG 1331
Qy 551 ArgPhePhe---------------ArgTyrLeuHisThrAlaValIleValSerGlyThr 565
Db 1332 GATGGGCTTCAGGGCCCAAGGGAGCGAGCCTTCCACACAGCCAGTGTTCTGGGCAATTAC 1391
Qy 566 MetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCys 585
Db 1392 ATGGTGGTCTATGGGGGCAATGTGCACACCCATTACCAGGAGGAA---------AAGTGC 1442
Qy 586 PheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrp--------------- 600
Db 1443 TACGAAGATGGCATCTTCTTCTACCACCTTGGCTGCCATCAATGGGTGTCAGGAGCTGAG 1502
Qy 601 ---------SerValLeuProArgProAspLeuHisHisAspValAsnArgPheGlyHis 617
Db 1503 CTTGCCCCGCCAGGAACCCCTGAGGGCCGAGCAGCGCCTCCCAGTGGTCGGTACTCACAT 1562
Qy 618 ---SerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeu 636
Db 1563 GTAGCTGCGGTGCTTGGTGGCAGCGTCCTGTTGGTGGCTGGGGGGTACAGCGGCCGGCCC 1622
Qy 637 LeuSerAspIleLeu---------------------ValPheThrSerGluGlnCysAsp 649
Db 1623 CGTGGGGACTTGATGGCGTACAAGGTGCCCCCCTTTGTGTTCCAGGCACCTGCCCCTGAC 1682
Qy 650 AlaHisArg-SerGluAlaAlaCys---------LeuAlaAlaGlyProGlyIleArg-- 665
Db 1683 TACCACTTGGACTACTGCTCCATGTACACAGACCACAGCGTCTGCTCCCGGGACCCGGAA 1742
Qy 666 CysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGln 685
Db 1743 TGCAGTTGGTGCCAAGGAGCCTGCCAA-----------------GCTGCACCCCCTCCT 1784
Qy 686 GluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArgCysAsp 705
Db 1785 GGGACCCCCTTG---GGGGCTTGTCCAGCCGCCAGCTGCCTGGGCCTGGGCCGCCTC--- 1838
Qy 706 GlnHisThrAspCysTyrSerCysThrAla--------------------------Asn 716
Db 1839 ---CTGGGTGACTGCCAGGCCTGCCTGGCCTTCAGCAGCCCCACAGCCCCTCCACGGGGA 1895
Qy 717 ThrAsnAspCysHisTrpCys------AsnAspHisCysValProArgAsnHisSer--- 733
Db 1896 CCTGGCACCCTGGGCTGGTGCGTGCACAATGAGAGCTGCCTCCCTAGGCCTGAGCAGGCC 1955
Qy 734 ---CysSerGluGlyGlnIleSer--------------------------IlePhe 742
Db 1956 CGCTGCCGAGGGGAGCAGATCTCAGGCACTGTGGGCTGGTGGGGGCCTGCGCCTGTCTTC 2015
Qy 743 ArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCys-AsnLysLysThrSerCy 762
Db 2016 --------------------GTCACGTCCCTGGAGGCCTGCGTCACCCAGAGCTTCCTG 2054
Qy 762 sArgSerCys-- 765
Db 2055 CCTGGCCTGCACTTGCTCACCTTTCAGCAGCCGCCCAATACCTCCCAGCCTGACAAGGTC 2114


Qy 766 ----------AlaLeuAspGlnAsnCysGlnTr 773
Db 2115 TCAATTGTCCGCAGCACGACCATCACCCTAACACCCAGCGCAGAGACAGATGTGTCCCTG 2174
Qy 773 pGluPro----------Ar 776
Db 2175 GTCTACCGTGGCTTCATCTACCCAATGCTGCCTGGAGGGCCAGGTGGACCAGGGGCTGAG 2234
Qy 776 gAsnGlnGluCys----------IleAl 782
Db 2235 GACGTGGCCGTGTGGACGCGGGCCCAGCGCCTACACGTCCTGGCCCGGATGGCCCGTGGC 2294
Qy 782 aLeuProGluAsnIle---------CysGlyIleGlyTrp---------- 792
Db 2295 CCTGACACGGAGAACATGGAGGAGGTGGGGCGCTGGGTGGCTCATCAGGAGAAGGAGACG 2354
Qy 793 ----------HisLeuValGlyAsnSerCysLeuLysIleThrThrAl 805
Db 2355 CGGCGGCTGCAGCGCCCTGGGTCTGCTCGCCTCTTCCCTCTGCCTGGGCGGGACCACAAG 2414
Qy 805 aLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAla-- 824
Db 2415 TATGCAGTAGAGATCCAGGGCCAGCTCAATGGCTCGGCAGGCCCTGGGCACAGCGAGCTA 2474
Qy 824 ---------- 824
Db 2475 ACTCTGCTGTGGGATCGGACTGGTGTGCCAGGAGGCAGCGAGATCTCCTTCTTCTTCCTG 2534
Qy 825 -SerLeuThrThrGlnLysLysValGluPheValLeuLys----------GlnLe 839
Db 2535 GAGCCCTACCGCTCGTCGTCCTGCACCTCCTATTCTTCCTGCCTGGGCTGCTTGGCAGAC 2594
Qy 839 uArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThr---------ProTr 856
Db 2595 CAGGGCTGTGGCTGGTGCCTGACCAGTGCCACCTGCCACCTGCGCCAGGGCGGAGCCCAT 2654
Qy 856 pValGlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheTh 876
Db 2655 TGCGGGGATGACGGGGCTGGTGGGTCCCTGCTGGTGCTGG------TGCCTACCCTCTGC 2708
Qy 876 rAsnSerLeuLeuGlnTrp----------MetProSerGluProSerAspAlaGl 891
Db 2709 CCACTCTGCGAGGAGCATCGGGACTGCCACGCCTGCACCCAGGACCCCTTCTGTGAGTGG 2768
Qy 891 yPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThrCysIleAs 911
Db 2769 C----------ATCAGAGCACCAGCCGCAAAGGGGACGCGGCATGCAGCC-------- 2808
Qy 911 nProLeuAsnGlySerValCysGluArgProAlaAsnHisSerAlaLysGlnCysArgTh 931
Db 2809 ----------GGCGGGGCCGG-GGTCGGGGTGCCCTGAAGAGTCCAGAGGAGTGTCCCCC 2857
Qy 931 rProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTr 951
Db 2858 GCTCTGCAGCCAGCGACTGACCTGTGAGGACTGCCTGGCCAACTCTAGCCAGTGCGCCTG 2917
Qy 951 pCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrValAlaSerProProGl 971
Db 2918 GTGCCAGTCCACCCACACCTGCTTCCTGTTTGCTGCCTACTTGGCCCGGTACCCACACGG 2977
Qy 971 yGlnCysMetGluTrp----------TyrThrMetSerThrCysProProGluAsnCy 987
Db 2978 GGGCTGTCGAGGCTGGGACGACAGTGTACACTCGGAGCCACGGTGC------CGGAGCTG 3031
Qy 987 sSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCysThrAs 1007
Db 3032 CGATGGCTTCCTGACCTGCCATGAGTGTCTGCAGAGCCACGAGTGTGGCTGGTGTGGCAA 3091
Qy 1007 pProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyPro-------- 1024
Db 3092 TGAGGACAACCCCACACTGGGACGGTGCCTACAGGGGGACTTCTCAGGGCCCCTCGGTGG 3151
Qy 1025 ----------ValLysMetProSerGlnAlaProThrGl 1034

```
Db       3152 GGGTAACTGCTCCCTGTGGGTGGGGGAGGGCCTGGGGCTTCCCGTGGCCCTCCCTGCCCG 3211

Qy       1034 yAsnPheTyrProGln---------------------------------------------- 1039

Db       3212 CTGGGCATACGCCCGCTGTCCTGACGTGGATGAGTGTCGCCTGGGCCTGGCCCGGTGCCA 3271

Qy       1040 ----------ProLeuLeuAsnSerSerMet---------------------------- 1046

Db       3272 CCCGCGGGCGACCTGCCTGAACACGCCCCTCAGCTACGAGTGTCACTGCCAGCGGGGCTA 3331

Qy       1046 ------------------------------------------------------------ 1046

Db       3332 CCAGGGTGATGGCATCTCACACTGCAACCGCACGGATGGCATCTCACACTGCAACCGCAC 3391

Qy       1047 -CysLeuGluAspSerArgTyr---------------------------------------- 1053

Db       3392 GTGCTTGGAGGACTGTGGCCATGGTGTGTGCAGTGGCCCCCCGGACTTTACCTGCGTGTG 3451

Qy       1054 -------AsnTrpSerPheIleHisCysProAla---------------------------- 1062

Db       3452 TGACCTAGGCTGGACATCAGACCTGCCCCCTCCCACACCCGCCCCGGGTCCGCCAGCCCC 3511

Qy       1063 ----------------CysGlnCysAsnGlyHisSerLysCysIleAsnGln------Se 1075

Db       3512 CCGCTGCTCCCGGGACTGTGGCTGCAGCTTCCACAGCCACTGCCGCAAGCGGGGCCCTGG 3571

Qy       1075 rIleCysGluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGl 1095

Db       3572 CTTCTGCGACGAGTGCCAGGACTGGACATGGGGGGAGCACTGCGAACGATGCCGGCCCGG 3631

Qy       1095 yPheTyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAl 1115

Db       3632 CAGCTTCGGCAACGCCACAGGCTCTAGGGGCTGCCGGCCCTGCCAGTGCAACGGGCACGG 3691

Qy       1115 a------------SerLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGl 1131

Db       3692 GGACCCACGCCGTGGCCACTGCGACAACCTCAGTGGGCTCTGCTTCTGCCAGGACCAC-- 3749

Qy       1131 yValLysGlyAspGluCysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnPro-- 1150

Db       3750 -ACCGAGGGTGCCCACTGCCAGCTCTGC------TCCCCAGGCTATTATGGGGATCCCCG 3802

Qy       1151 -LeuArgGlyThrCysTyr 1156

Db       3803 GGCCGGTGGTTCCTGCTTT 3821

RESULT 10
US-10-120-988-143
; Sequence 143, Application US/10120988
; Publication No. US20030219745A1
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Goodrich, Ryle
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. US20030219745A1el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 802CON
;  CURRENT APPLICATION NUMBER: US/10/120,988
;  CURRENT FILING DATE:  2002-04-11
;  PRIOR APPLICATION NUMBER: 09/774,528
;  PRIOR FILING DATE: 2001-01-30
;  NUMBER OF SEQ ID NOS: 441
;  SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 143
;   LENGTH: 7244
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (714)..(6518)
US-10-120-988-143

Alignment Scores:
Pred. No.:              3.69e-79        Length:         7244
Score:                  923.00          Matches:        423
Percent Similarity:     32.95%          Conservative:   206
Best Local Similarity:  22.16%          Mismatches:     576
Query Match:            11.66%          Indels:         706
DB:                     15              Gaps:           92

US-09-787-097-12 (1-1429) x US-10-120-988-143 (1-7244)

Qy          2 ValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAla-Al 21

Db       1279 GTCACTGCCAGCGGGGCTACC------AGGGTGATGGCATCTCACACTGCAACCGCACGT 1332

Qy         21 aLeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArgPr 41

Db       1333 GCTTGGAGGACTG--------------TGGCCATGGTGTGTG------CAGTGGCCCCC 1371

Qy         41 oGlyLeuGlyAlaGlyLeuArgLeu---ProArgLeu---------------------- 52

Db       1372 CGGACTT------TACCTGCGTGTGTGACCTAGGCTGGACATCAGACCTGCCCCCTCCCA 1425

Qy         53 ----------------LeuSerProProLeuArgProArgLeuLeuLeuLeuLeuLeLe 67

Db       1426 CACCCGCCCCGGGTCCGCCAGCCCCCCGCTGCTCCCGGGA---------CTGTGGCTGCA 1476

Qy         67 uLeuProProProLeu-----------------LeuLeuLeuLeuLeuProCysGluAl 81

Db       1477 GCTTCCACAGCCACTGCCGCAAGCGGGGCCCTGGCTTCTGCGACGAGTGCCAGG-ACTGG 1535

Qy         81 aGluAlaAlaAlaAlaAla------AlaAlaValSerGlySerAlaAlaAlaGluAlaLy 99

Db       1536 ACATGGGGGGAGCACTGCGAACGATGCCGGCCCGGCAGCTTCGGCAACGCCACAGGCTCT 1595

Qy         99 sGluCysAspArgProCys---ValAsnGly-----------------GlyArgCysAs 112

Db       1596 AGGGGCTGC-CGGCCCTGCCAGTGCAACGGGCACGGGGACCCACGCCGTGGCCACTGCGA 1654

Qy        112 nProGlyThrGlyGlnCysValCysProAlaGlyTrpValGlyGluGlnCysGln----- 130

Db       1655 CAACCTCAGTGGGCTCTGCTTCTGCCAGGACCACACCGAGGGTGCCCACTGCCAGCTCTG 1714

Qy        131 ----------------------------------------------HisCysGlyGl 134

Db       1715 CTCCCCAGGCTATTATGGGGATCCCCGGGCCGGTGGTTCCTGCTTTCGGGAGTGTGGGGG 1774

Qy        134 yArgPheArgLeuThrGlySerSer---------------------------------- 142

Db       1775 TCGCGCCCTCCTCACCAACGTGTCCTCAGTGGCACTGGGCTCACGCCGGGTCGGGGGGCT 1834

Qy        143 ----------------GlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLy 157

Db       1835 GCTGCCTCCAGGTGGCGGGGCTGCAAGAGCCGGGCCTGGC------------CTGTCCTA 1882

Qy        157 sCysThrTrpLeuIleGluGly--------------------GlnProAsnArgIle-- 169

Db       1883 CTGTGTGTGGGTTGTCTCGGCCACTGAGGAGCTACAGCCCTGTGCTCCCGGGACCCTCTG 1942

Qy        170 -------MetArgLeuArgPheAsn---HisPheAlaThrGluCysSerTrpAspHisLe 186

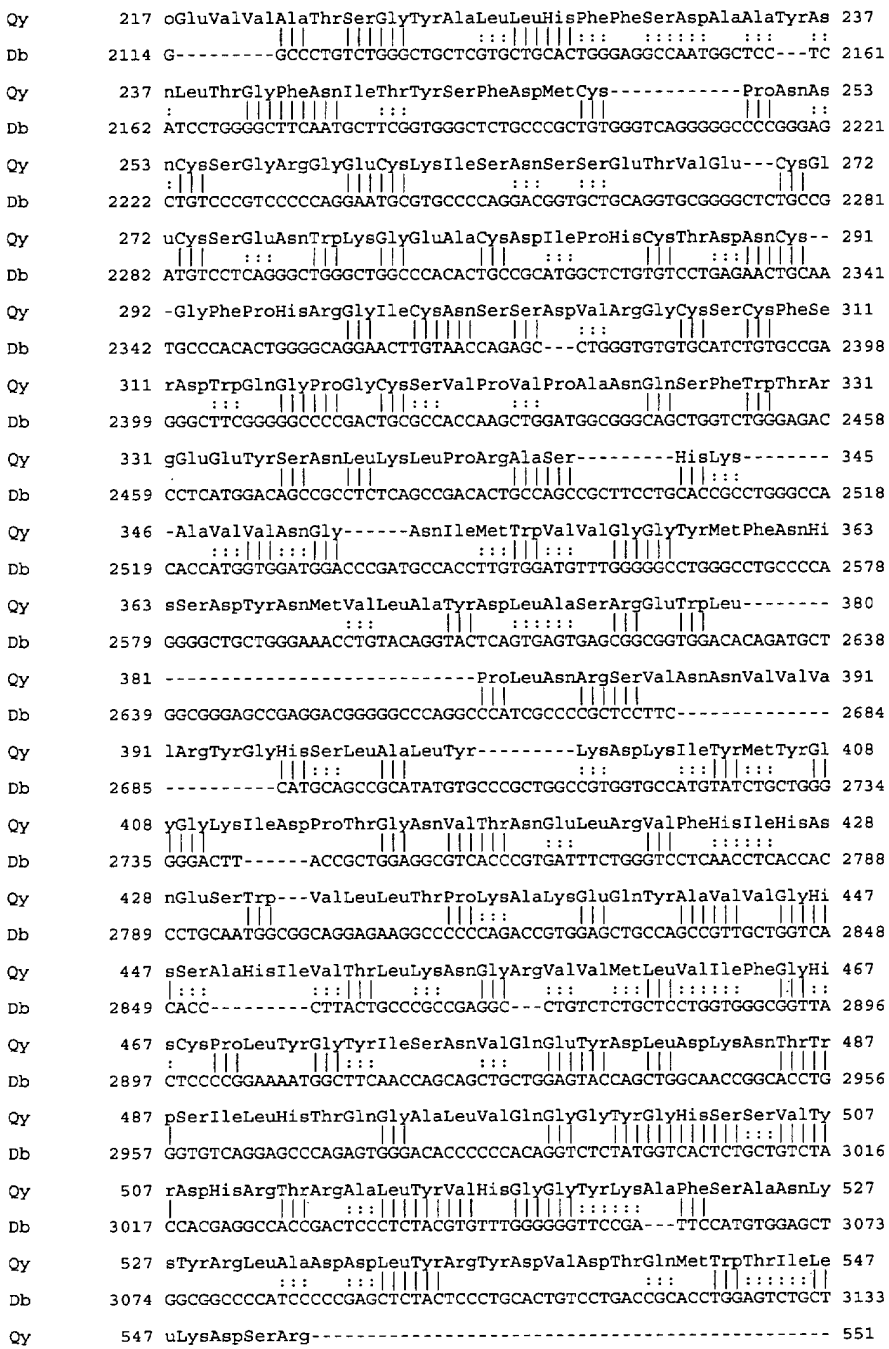
Db       1943 TCCCCCACTCACCCTCACCTTCTCCCCCGACAGCAGCACCCCCTGCACGCTGAGCTACGT 2002

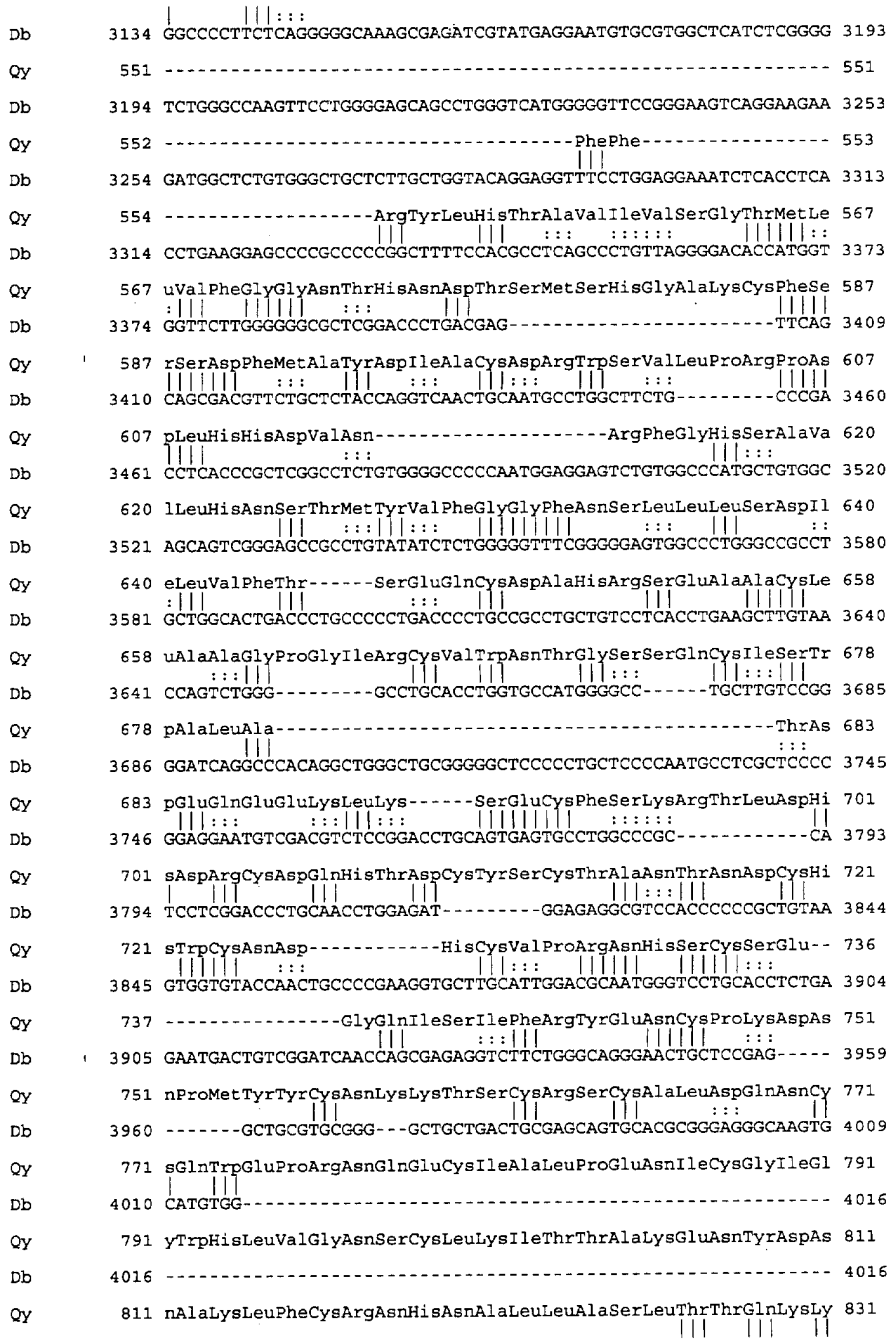
Qy        186 uTyrValTyrAspGlyAspSerIleTyr-------------------------AlaPr 197

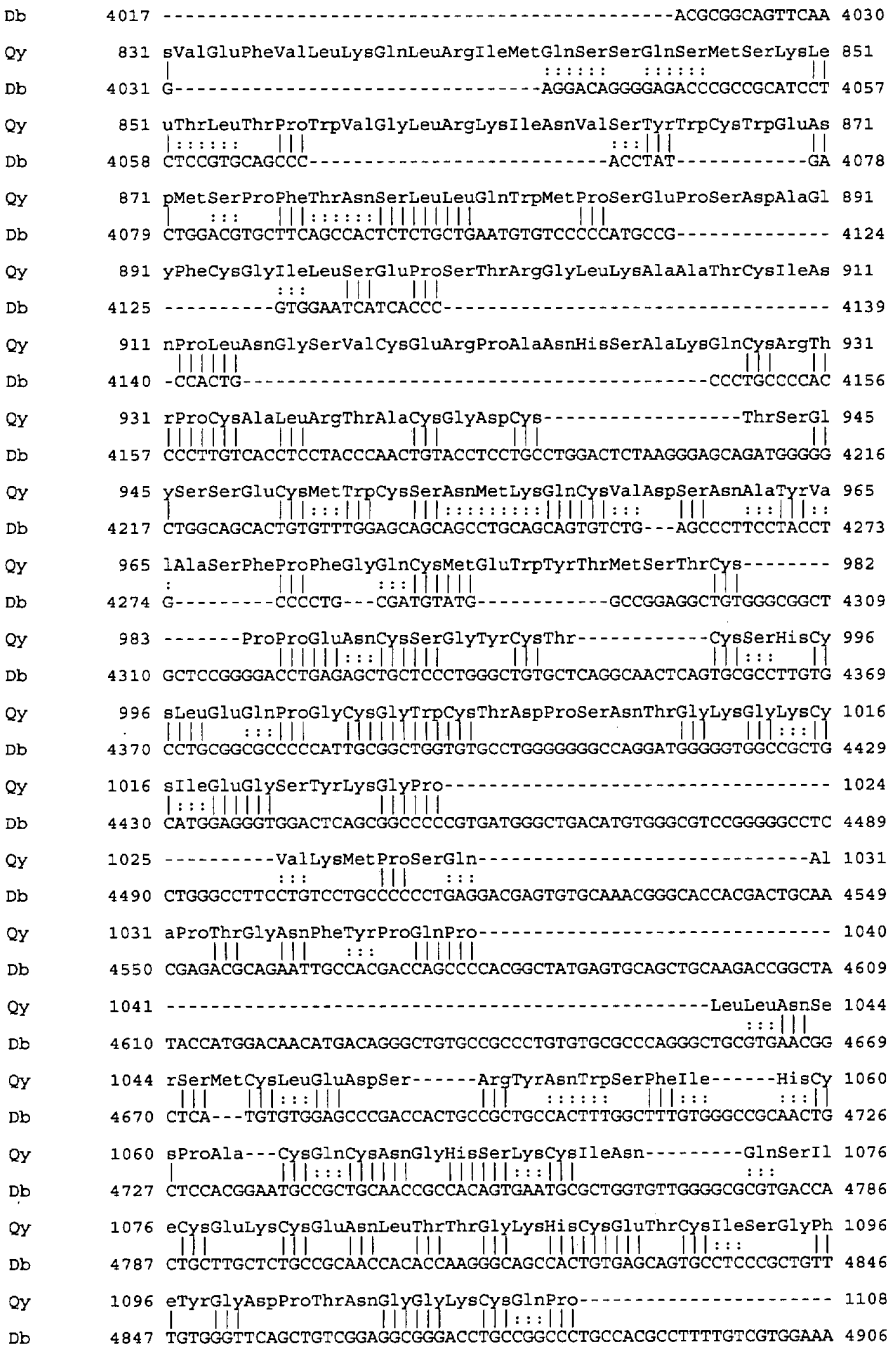
Db       2003 CCTGGCGTTTGATGGATTCCCACGCTTCCTGGACACTGGTGTTGTCCAGTCGGACCGCAG 2062

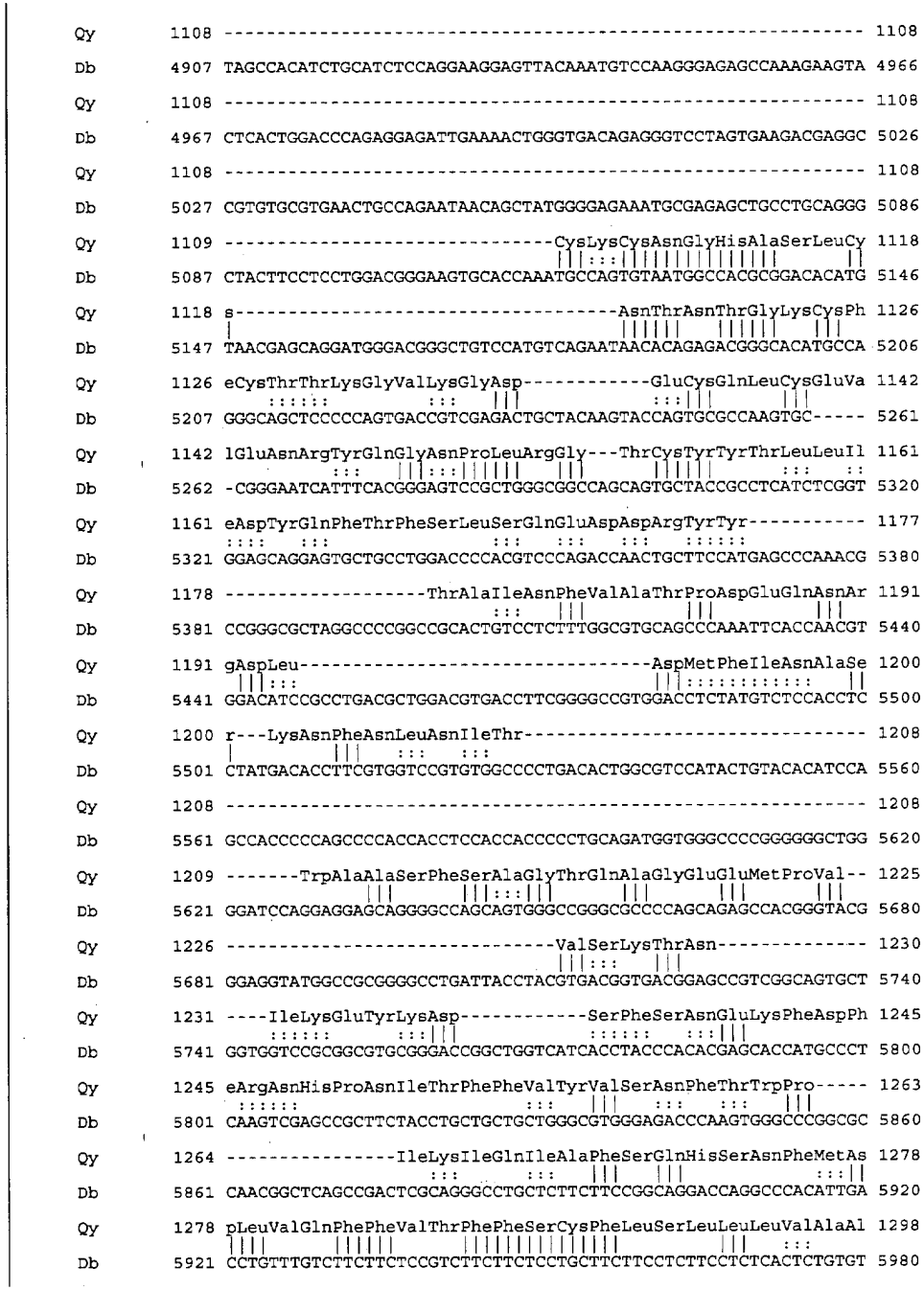
Qy        197 oLeuValAlaAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValPr 217

Db       2063 CCTCATAGCTGCCTTCTGCGGC---------CAGCGACGGGACAGGCCCCTCACTGTTCA 2113
```









```
Qy       1298 aValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeuLeuAr 1318
Db       5981 GCTCCTCTGGAAGGCCAAGCAGGCTCTGGACCAGCGGCAGGAGCAGCGCCGGCACTTGCA 6040

Qy       1318 gGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThrAs 1338
Db       6041 GGAGATGACCAAGATGGCCAGCCGCCCCTTCGCCAAGGTCACCGTCTGCTTCCCACCTGA 6100

Qy       1338 p------------------------------------------------------------ 1338
Db       6101 CCCTACTGCCCCGGCCTCCGCCTGGAAGCCGGCTGGGCTCCCACCTCCCGCCTTCCGCCG 6160

Qy       1339 -GluGluPro---------ProAspLeuIleGlyGlySer-------------------- 1348
Db       6161 CTCTGAGCCCTTCCTGGCACCCCTGCTGCTGACAGGGGCCGGTGGGCCCTGGGGACCCAT 6220

Qy       1349 ----------------IleLysThrValProLys-----------------------Pr 1355
Db       6221 GGGAGGGGGCTGCTGCCCACCAGCCATCCCCGCCACCACTGCTGGGCTGCGAGCTGGGCC 6280

Qy       1355 oIleAlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLe 1375
Db       6281 CATCACTCTCGAGCCCACAGAAGATGGCATGGCTGGCGTGGCCACACTGCTGCTCCAGCT 6340

Qy       1375 uProArgGlyLeuGlyGlyIleProProProGlyGlnSerGlyLeuAlaValAlaSerAl 1395
Db       6341 GCCTGGCGGG--------------CCCCATGCACCCAACGGCGCCTGCCTGGGGTCAGC 6385

Qy       1395 aLeuValAspIleSerGlnGlnMet 1403
Db       6386 CCTCGTCACACTGCGGCACAGGCTG 6410

RESULT 11
US-10-277-216-5
; Sequence 5, Application US/10277216
; Publication No. US20040002470A1
; GENERAL INFORMATION:
;  APPLICANT: KEITH, TIM
;  TITLE OF INVENTION: NOVEL HUMAN GENE RELATING TO RESPIRATORY DISEASES,
;  TITLE OF INVENTION:  OBESITY, AND INFLAMMATORY BOWEL DISEASE
;  FILE REFERENCE: 2976-4051
;  CURRENT APPLICATION NUMBER: US/10/277,216
;  CURRENT FILING DATE:  2002-10-17
;  PRIOR APPLICATION NUMBER: 10/126,022
;  PRIOR FILING DATE: 2002-04-19
;  PRIOR APPLICATION NUMBER: 09/834,597
;  PRIOR FILING DATE: 2001-04-13
;  PRIOR APPLICATION NUMBER: 09/548,797
;  PRIOR FILING DATE: 2000-04-13
;  NUMBER OF SEQ ID NOS: 420
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 207433
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-277-216-5

Alignment Scores:
Pred. No.:              1.3e-68       Length:         207433
Score:                  840.00        Matches:        231
Percent Similarity:     24.34%        Conservative:   10
Best Local Similarity:  23.33%        Mismatches:     34
Query Match:            10.61%        Indels:         717
DB:                     15            Gaps:           6

US-09-787-097-12 (1-1429) x US-10-277-216-5 (1-207433)

Qy        796 GlyAsnSerCysLeuLysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPhe 815
Db       7134 GGTGGCTCATGCCTATA-ATCTCAACATTTTGGGAGGCCAAGGTGGGAGGATCACTTGAG 7192

Qy        816 CysArgAsnHisAsnAlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheVal 835
```

```
Db       7193 CTCAGG------AGTTTGAAACTAGCCAGGGCAACATGGCAAGACTCCATCTCCAAAATT 7246

Qy        836 ---LeuLysGlnLeuArgIleMetGlnSerSerGlnSerMet-----------SerLys 850
Db       7247 TAAATTAAAAAAAGATACTATCTGTATTCATAGTTGTGTCTCTTTTGCCTTTAGTCCAAG 7306

Qy        851 LeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCysTrpGlu 870
Db       7307 CTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAA 7366

Qy        871 AspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSerAspAla 890
Db       7367 GATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCT 7426

Qy        891 GlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThrCysIle 910
Db       7427 GGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATC 7486

Qy        911 AsnProLeuAsnGlySerValCysGluArg---------------------------- 920
Db       7487 AACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGGTAAGTTCACAGGTGAATTAGGTGGT 7546

Qy        920 ------------------------------------------------------------ 920
Db       7547 ATTCAGAGTTTATTGTGAGAGAAACCATAGGAGGCATAGTTCATTGCTGAGATGTGTGAA 7606

Qy        920 ------------------------------------------------------------ 920
Db       7607 GTAGTCATGAAAACAGATGAAGTATTGATTTCAAGCATGCAAAGAAGAGTATAACCCAGA 7666

Qy        920 ------------------------------------------------------------ 920
Db       7667 TTTCAGAAGCAGAAGGAAATATTCTGGGACCCTGAATAGTTTTAATTATAAGCAAAACTA 7726

Qy        920 ------------------------------------------------------------ 920
Db       7727 AAAATAACTAACACTACTCGAAGAAACTGATATTCTAATTAACAATGAGATTGATAGGTT 7786

Qy        920 ------------------------------------------------------------ 920
Db       7787 TATTGACCAGAAAAAGTATTGAGAATTGCTCTGAAAAGCAAATTTATTGGTGTTGGCAGA 7846

Qy        920 ------------------------------------------------------------ 920
Db       7847 GAAATGCTGTGGAAGAAGAAACAAAAAAGGAAAATAAAACCAAAGAAAAGATTAGTAAAG 7906

Qy        920 ------------------------------------------------------------ 920
Db       7907 CAAGCAAGTGACTGCAGGGACAGTGTTCAGAAAGGTAGTGTCAACAGGGAGAAAAATGAT 7966

Qy        920 ------------------------------------------------------------ 920
Db       7967 GAGAGCAGTTCTGTAAGCGAGGACAGAAGCAACCCAGAACTACGCGAGAGCTGCAGGAGA 8026

Qy        920 ------------------------------------------------------------ 920
Db       8027 GTACTGAGCAGACAGACACTCGGAGTAGCTTCCCAGCTTTCAGTCTCTCTGGGCTACATT 8086

Qy        920 ------------------------------------------------------------ 920
Db       8087 TTGGCTAACTAAAGGCAGGCCCAGGGGTAGCTGCTGCAGCATGAAGACAACTCAAAATAA 8146

Qy        920 ------------------------------------------------------------ 920
Db       8147 GCCTCCTCCTCCTGCAGAGGGACTCACAAGCAGTCGGTGGAATTGTGGGTTATTTTGGGA 8206

Qy        920 ------------------------------------------------------------ 920
Db       8207 GGCTGGACTCCATTTGCATTGTGTCCAATTTTGGGAAAGTAATTTGTCTGAGGAATTACA 8266

Qy        920 ------------------------------------------------------------ 920
Db       8267 GAGGTATAGGAGAGAATTCATAGCATTGTAGATTTCTAAATATTGATTTCTAAACTTCTC 8326
```

```
Qy       920 ------------------------------------------------------------ 920
Db      8327 TAGGGGAGCTGCCCAGCAGCCTTTCAGAATTCTGAATCCTTTGTTGAACTTAATGAATGC 8386
Qy       920 ------------------------------------------------------------ 920
Db      8387 CGTAGACCTTCTTCCCTCCAAAATTGCAAACATGAGATTTTACATATTCAAGTGGATTGT 8446
Qy       920 ------------------------------------------------------------ 920
Db      8447 AAAACCCCTAAAGTTCATCCTGGGTCCCCAAATTCTAAGGGGCTTACAGCCCTACTTTTA 8506
Qy       920 ------------------------------------------------------------ 920
Db      8507 TGGAAAGATTCACTGTTTCCTCATCTTTGTCCTTAATGCTCTTTGAAAAGAAAATTTATT 8566
Qy       920 ------------------------------------------------------------ 920
Db      8567 TTTTCCACAAGTGTGTAAATACACTTGACTAAACAGTACTTGATGACTTTTATTGTTTTC 8626
Qy       920 ------------------------------------------------------------ 920
Db      8627 TATTTTCTCTCATTTAATTGATTTCATGGCACATTAATGGGTCATCACCCACATTTGAAA 8686
Qy       920 ------------------------------------------------------------ 920
Db      8687 AGTCTTGGCTGGGCACAGTGGCTCACGCCTGTAATCCCAGCACTTTAGGAGGCCAAGACG 8746
Qy       920 ------------------------------------------------------------ 920
Db      8747 GGTGGATCACAAGGTCAAGATATCAAGACCATCCTGGCTAACATAGTGAAACCCCATCTC 8806
Qy       920 ------------------------------------------------------------ 920
Db      8807 TACTAAAAATACAAAAAATTAGCCAGACGTGATGGCGCACCCCTGTAGTCCCAGCTACTT 8866
Qy       920 ------------------------------------------------------------ 920
Db      8867 GGGAGGCTGAGGCAGGAGAATCACTTGAACCGGGGAGGCGGAGGTTGCAGTGAGCCGAAA 8926
Qy       920 ------------------------------------------------------------ 920
Db      8927 TTGTGCCACTGCACTCTAGCCTGGAGACAGAGTGAGACTCCGTCTCAAAAAAAAAAAAGA 8986
Qy       920 ------------------------------------------------------------ 920
Db      8987 AAAGAAAGAAAGAAAAGTCTTTGACCTTAGCGGACATGGTGGGAGCTCTAAGTGTCTCTC 9046
Qy       921 ------------ProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuAr 936
                          ||||||||||||||||||||||||||||||||||||||||||||
Db      9047 TTGGGTTTCATTCCCAGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAG 9106
Qy       936 gThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLy 956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9107 GACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAA 9166
Qy       956 sGlnCysValAspSerAsnAlaTyrValAlaSerProProGlyGlnCysMetGluTr 976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9167 GCAGTGTGTGGACTCCAATGCCTACGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATG 9226
Qy       976 pTyrThrMetSerThrCysProPro---------------------------------- 984
             |||||||||||||||||   |||
Db      9227 GTATACGATGAGCACCTGCCC-CCGTAAGTGAAAAAGGGAGCCCTAGGCACTTATGCATG 9285
Qy       984 ------------------------------------------------------------ 984
Db      9286 CCCTCTGTATAGGCAACAACTCAGCCATGAGGCTGTGCTGTCAGCCTCTGAACATTTTAG 9345
Qy       984 ------------------------------------------------------------ 984
Db      9346 AAACAAGACTGGACATGACCTCTGCTCAAACCTGACCAGAGACTGCCATCGAGACCTTGC 9405
```

```
Qy       984 ------------------------------------------------------------ 984
Db      9406 TGCCTATTGAGAACCTTCATACAGAATCAGGCACATTGACAGTAAATAAATGTAAGATAG 9465
Qy       984 ------------------------------------------------------------ 984
Db      9466 ATCACAGAGTACAGAAATAACTTGTCCAACTTCAGTGTCATATTGCTCAATCCATGTAAT 9525
Qy       984 ------------------------------------------------------------ 984
Db      9526 ATCTCCATATCTGAATTTCCTAATTTGTAAAGTAAATGCTTTCCAATAGATAATCTCTAA 9585
Qy       984 ------------------------------------------------------------ 984
Db      9586 GGTCCCTTTTGCCTTCAACATCCTGGGATTGAGAGAGGAGGGAAGGGTCATCTCTGTTAT 9645
Qy       984 ------------------------------------------------------------ 984
Db      9646 GTATTGGGCAAAATACTGGGCTCTTTACATTCATTATCTCTTTTAAATAATCAAGACAGA 9705
Qy       984 ------------------------------------------------------------ 984
Db      9706 ATAATATTTTTGACTCAAGCCAGTTGAATAGTCTGTTAAAAAAAAAGTAAATACAGTGAA 9765
Qy       985 ----------------------------------------------------GluAsnCysSe 988
                                                                 ||||||||||
Db      9766 TTCAGATCTACCTGTGATAGTCAATTGCAACTTTTTTTTTTTAATAGCTGAAAATTGTTC 9825
Qy       988 rGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCysThrAspPr 1008
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9826 AGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCC 9885
Qy      1008 oSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyProValLysMetPr 1028
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9886 CAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCC 9945
Qy      1028 oSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerSerMetCysLe 1048
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9946 TTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCT 10005
Qy      1048 uGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnCysAsnGlyHi 1068
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     10006 AGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCA--------------GGTAA 10050
Qy      1068 sSerLysCysIleAsnGlnSerIleCys 1077
                    ||||||::::::      |||
Db     10051 GATGCCTTGCATATCCAAATTCAAGTGT 10078
```

RESULT 12
US-10-126-022-5
; Sequence 5, Application US/10126022
; Publication No. US20040023215A1
; GENERAL INFORMATION:
; APPLICANT: KEITH, TIM
; TITLE OF INVENTION: NOVEL HUMAN GENE RELATING TO RESPIRATORY DISEASES,
; TITLE OF INVENTION: OBESITY, AND INFLAMMATORY BOWEL DISEASE
; FILE REFERENCE: 2976-4039US2
; CURRENT APPLICATION NUMBER: US/10/126,022
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 09/834,597
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 09/548,797
; PRIOR FILING DATE: 2000-04-13
; NUMBER OF SEQ ID NOS: 420
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 207433
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-126-022-5

Alignment Scores:
Pred. No.: 1.3e-68 Length: 207433

Score: 840.00 Matches: 231
Percent Similarity: 24.34% Conservative: 10
Best Local Similarity: 23.33% Mismatches: 34
Query Match: 10.61% Indels: 717
DB: 16 Gaps: 6

US-09-787-097-12 (1-1429) x US-10-126-022-5 (1-207433)

```
Qy       796 GlyAsnSerCysLeuLysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPhe 815
Db      7134 GGTGGCTCATGCCTATA-ATCTCAACATTTTGGGAGGCCAAGGTGGGAGGATCACTTGAG 7192

Qy       816 CysArgAsnHisAsnAlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheVal 835
Db      7193 CTCAGG------AGTTTGAAACTAGCCAGGGCAACATGGCAAGACTCCATCTCCAAAATT 7246

Qy       836 ---LeuLysGlnLeuArgIleMetGlnSerSerGlnSerMet------------SerLys 850
Db      7247 TAAATTAAAAAAAGATACTATCTGTATTCATAGTTGTGTCTCTTTTGCCTTTAGTCCAAG 7306

Qy       851 LeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCysTrpGlu 870
Db      7307 CTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAA 7366

Qy       871 AspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSerAspAla 890
Db      7367 GATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCT 7426

Qy       891 GlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThrCysIle 910
Db      7427 GGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATC 7486

Qy       911 AsnProLeuAsnGlySerValCysGluArg---------------------------- 920
Db      7487 AACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGGTAAGTTCACAGGTGAATTAGGTGGT 7546

Qy       920 ------------------------------------------------------------ 920
Db      7547 ATTCAGAGTTTATTGTGAGAGAAACCATAGGAGGCATAGTTCATTGCTGAGATGTGTGAA 7606

Qy       920 ------------------------------------------------------------ 920
Db      7607 GTAGTCATGAAAACAGATGAAGTATTGATTTCAAGCATGCAAAGAAGAGTATAACCCAGA 7666

Qy       920 ------------------------------------------------------------ 920
Db      7667 TTTCAGAAGCAGAAGGAAATATTCTGGGACCCTGAATAGTTTTAATTATAAGCAAAACTA 7726

Qy       920 ------------------------------------------------------------ 920
Db      7727 AAAATAACTAACACTACTCGAAGAAACTGATATTCTAATTAACAATGAGATTGATAGGTT 7786

Qy       920 ------------------------------------------------------------ 920
Db      7787 TATTGACCAGAAAAAGTATTGAGAATTGCTCTGAAAAGCAAATTTATTGGTGTTGGCAGA 7846

Qy       920 ------------------------------------------------------------ 920
Db      7847 GAAATGCTGTGGAAGAAGAAACAAAAAAGGAAAATAAAACCAAAGAAAAGATTAGTAAAG 7906

Qy       920 ------------------------------------------------------------ 920
Db      7907 CAAGCAAGTGACTGCAGGGACAGTGTTCAGAAAGGTAGTGTCAACAGGGAGAAAAATGAT 7966

Qy       920 ------------------------------------------------------------ 920
Db      7967 GAGAGCAGTTCTGTAAGCGAGGACAGAAGCAACCCAGAACTACGCGAGAGCTGCAGGAGA 8026

Qy       920 ------------------------------------------------------------ 920
Db      8027 GTACTGAGCAGACAGACACTCGGAGTAGCTTCCCAGCTTTCAGTCTCTCTGGGCTACATT 8086

Qy       920 ------------------------------------------------------------ 920
Db      8087 TTGGCTAACTAAAGGCAGGCCCAGGGGTAGCTGCTGCAGCATGAAGACAACTCAAAATAA 8146

Qy       920 ------------------------------------------------------------ 920
Db      8147 GCCTCCTCCTCCTGCAGAGGGACTCACAAGCAGTCGGTGGAATTGTGGGTTATTTTGGGA 8206

Qy       920 ------------------------------------------------------------ 920
Db      8207 GGCTGGACTCCATTTGCATTGTGTCCAATTTTGGGAAAGTAATTTGTCTGAGGAATTACA 8266

Qy       920 ------------------------------------------------------------ 920
Db      8267 GAGGTATAGGAGAGAATTCATAGCATTGTAGATTTCTAAATATTGATTTCTAAACTTCTC 8326

Qy       920 ------------------------------------------------------------ 920
Db      8327 TAGGGGAGCTGCCCAGCAGCCTTTCAGAATTCTGAATCCTTTGTTGAACTTAATGAATGC 8386

Qy       920 ------------------------------------------------------------ 920
Db      8387 CGTAGACCTTCTTCCCTCCAAAATTGCAAACATGAGATTTTACATATTCAAGTGGATTGT 8446

Qy       920 ------------------------------------------------------------ 920
Db      8447 AAAACCCCTAAAGTTCATCCTGGGTCCCCAAATTCTAAGGGGCTTACAGCCCTACTTTTA 8506

Qy       920 ------------------------------------------------------------ 920
Db      8507 TGGAAAGATTCACTGTTTCCTCATCTTTGTCCTTAATGCTCTTTGAAAAGAAAATTTATT 8566

Qy       920 ------------------------------------------------------------ 920
Db      8567 TTTTCCACAAGTGTGTAAATACACTTGACTAAACAGTACTTGATGACTTTTATTGTTTTC 8626

Qy       920 ------------------------------------------------------------ 920
Db      8627 TATTTTCTCTCATTTAATTGATTTCATGGCACATTAATGGGTCATCACCCACATTTGAAA 8686

Qy       920 ------------------------------------------------------------ 920
Db      8687 AGTCTTGGCTGGGCACAGTGGCTCACGCCTGTAATCCCAGCACTTTAGGAGGCCAAGACG 8746

Qy       920 ------------------------------------------------------------ 920
Db      8747 GGTGGATCACAAGGTCAAGATATCAAGACCATCCTGGCTAACATAGTGAAACCCCATCTC 8806

Qy       920 ------------------------------------------------------------ 920
Db      8807 TACTAAAAATACAAAAAATTAGCCAGACGTGATGGCGCACCCCTGTAGTCCCAGCTACTT 8866

Qy       920 ------------------------------------------------------------ 920
Db      8867 GGGAGGCTGAGGCAGGAGAATCACTTGAACCGGGGAGGCGGAGGTTGCAGTGAGCCGAAA 8926

Qy       920 ------------------------------------------------------------ 920
Db      8927 TTGTGCCACTGCACTCTAGCCTGGAGACAGAGTGAGACTCCGTCTCAAAAAAAAAAAAGA 8986

Qy       920 ------------------------------------------------------------ 920
Db      8987 AAAGAAAGAAAGAAAAGTCTTTGACCTTAGCGGACATGGTGGGAGCTCTAAGTGTCTCTC 9046

Qy       921 ----------ProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuAr 936
Db      9047 TTGGGTTTCATTCCCAGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAG 9106

Qy       936 gThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLy 956
Db      9107 GACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAA 9166

Qy       956 sGlnCysValAspSerAsnAlaTyrValAlaSerProPheProGlyGlnCysMetGluTr 976
```

```
Db       9167 GCAGTGTGTGGACTCCAATGCCTACGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATG 9226

Qy        976 pTyrThrMetSerThrCysProPro---------------------------------- 984
              ||||||||||||||||||||    |||
Db       9227 GTATACGATGAGCACCTGCCC-CCGTAAGTGAAAAAGGGAGCCCTAGGCACTTATGCATG 9285

Qy        984 ---------------------------------------------------------- 984

Db       9286 CCCTCTGTATAGGCAACAACTCAGCCATGAGGCTGTGCTGTCAGCCTCTGAACATTTTAG 9345

Qy        984 ---------------------------------------------------------- 984

Db       9346 AAACAAGACTGGACATGACCTCTGCTCAAACCTGACCAGAGACTGCCATCGAGACCTTGC 9405

Qy        984 ---------------------------------------------------------- 984

Db       9406 TGCCTATTGAGAACCTTCATACAGAATCAGGCACATTGACAGTAAATAAATGTAAGATAG 9465

Qy        984 ---------------------------------------------------------- 984

Db       9466 ATCACAGAGTACAGAAATAACTTGTCCAACTTCAGTGTCATATTGCTCAATCCATGTAAT 9525

Qy        984 ---------------------------------------------------------- 984

Db       9526 ATCTCCATATCTGAATTTCCTAATTTGTAAAGTAAATGCTTTCCAATAGATAATCTCTAA 9585

Qy        984 ---------------------------------------------------------- 984

Db       9586 GGTCCCTTTTGCCTTCAACATCCTGGGATTGAGAGAGGAGGGAAGGGTCATCTCTGTTAT 9645

Qy        984 ---------------------------------------------------------- 984

Db       9646 GTATTGGGCAAAATACTGGGCTCTTTACATTCATTATCTCTTTTAAATAATCAAGACAGA 9705

Qy        984 ---------------------------------------------------------- 984

Db       9706 ATAATATTTTTGACTCAAGCCAGTTGAATAGTCTGTTAAAAAAAAAGTAAATACAGTGAA 9765

Qy        985 ----------------------------------------------------GluAsnCysSe 988
                                                                 ||||||||||
Db       9766 TTCAGATCTACCTGTGATAGTCAATTGCAACTTTTTTTTTTTAATAGCTGAAAATTGTTC 9825

Qy        988 rGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCysThrAspPr 1008
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9826 AGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCC 9885

Qy       1008 oSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyProValLysMetPr 1028
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9886 CAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCC 9945

Qy       1028 oSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerSerMetCysLe 1048
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9946 TTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCT 10005

Qy       1048 uGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnCysAsnGlyHi 1068
              ||||||||||||||||||||||||||||||||||||                  |||
Db      10006 AGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCA--------------GGTAA 10050

Qy       1068 sSerLysCysIleAsnGlnSerIleCys 1077
                  ||||||::::::      |||
Db      10051 GATGCCTTGCATATCCAAATTCAAGTGT 10078
```

RESULT 13
US-09-893-238-5
; Sequence 5, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 90050
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-893-238-5

Alignment Scores:

Pred. No.:	6.32e-62	Length:	90050
Score:	766.00	Matches:	205
Percent Similarity:	21.41%	Conservative:	2
Best Local Similarity:	21.20%	Mismatches:	6
Query Match:	9.68%	Indels:	755
DB:	9	Gaps:	2

US-09-787-097-12 (1-1429) x US-09-893-238-5 (1-90050)

```
Qy        849 SerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      28996 TCCAAGCTCACTCTGACTCCATGGGTTGGTCTTCGGAAGATCAATGTGTCTTACTGGTGC 29055

Qy        869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      29056 TGGGAGGATATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGT 29115

Qy        889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      29116 GATGCTGGCTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGATTAAAGGCTGCAACC 29175

Qy        909 CysIleAsnProLeuAsnGlySerValCysGluArgProAla------------------ 922
              ||||||||||||||||||||||||||||||||||||
Db      29176 TGCATCAACCCTCTCAATGGCAGCGTCTGTGAAAGGCCTGG-TAAGGACATGGGTGCATA 29234

Qy        922 ---------------------------------------------------------- 922

Db      29235 TAGTGCTCCAGGAGGAGCCAAGACAGCAAAGGAGGCACAGCTGAATGAGCGCTGAGGTGA 29294

Qy        922 ---------------------------------------------------------- 922

Db      29295 TGAAGTACTTATGGCAGCAGGGAGAGGAGCACCAATTTAGGCATATGTATTTCAAACAGA 29354

Qy        922 ---------------------------------------------------------- 922

Db      29355 ACCCGATTCCAGATAGTCTTTCTTGGCCTCTGACTGCTTTAAGCCATACTGAAAACCAAA 29414

Qy        922 ---------------------------------------------------------- 922

Db      29415 AATAAAATTGCTGAAAGAACCCAGTTTATATTGAGCTGCACTGTTTCGTTGGTCTCAAAG 29474

Qy        922 ---------------------------------------------------------- 922

Db      29475 TGTTGAGAATTGTTCTAGAAGATTATTTCCTTGGTGTTGGCAGAGAAGTGCTATGGAGGA 29534

Qy        922 ---------------------------------------------------------- 922

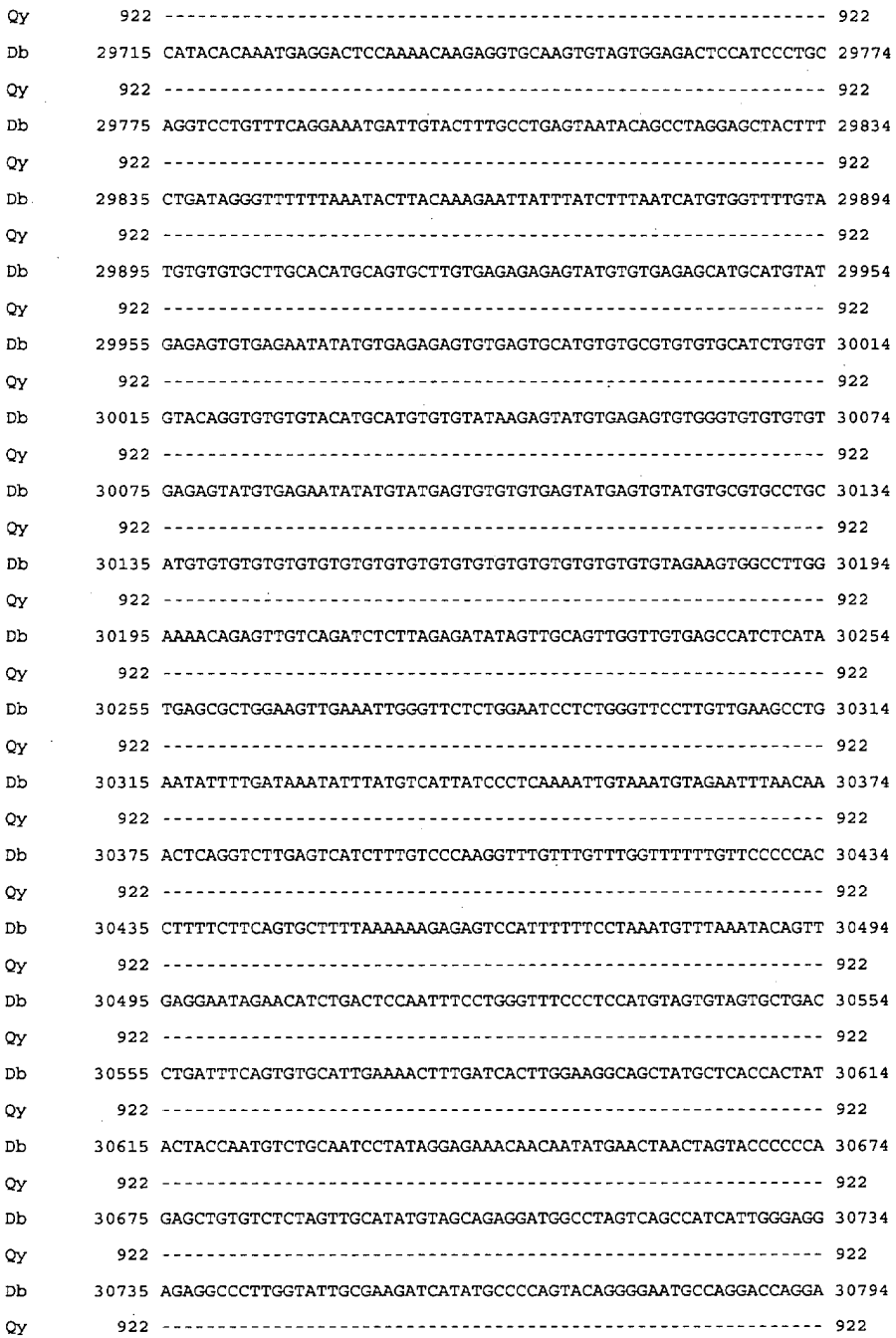
Db      29535 AACAACAACCTGAAACCAAAGAAACATTTAGAAAAGCAGCAAGTCAGGACACTATTCAGA 29594

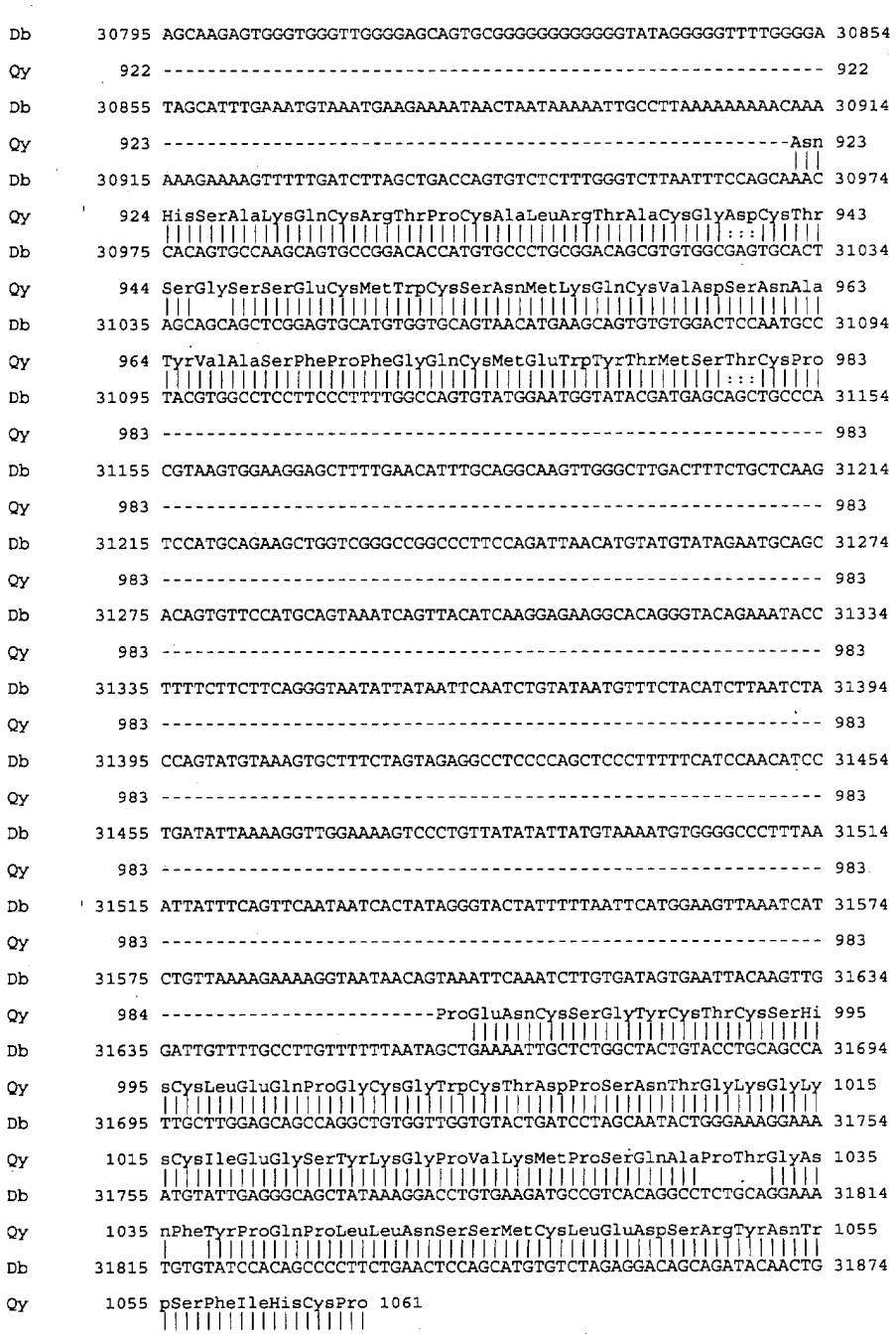
Qy        922 ---------------------------------------------------------- 922

Db      29595 CACTGCTGGGGTGGGGGGAGAGGGGCATGGCCAAAGAAGCCGACAGAGCCAACACCAGGC 29654

Qy        922 ---------------------------------------------------------- 922

Db      29655 TGTGGCAATGTCCTGCGCTGAGGTTAAGGTTAGACTCCATGAGGCCAGGCCCAGAACAGC 29714
```


Qy 922 ---------- 922
Db 29715 CATACACAAATGAGGACTCCAAAACAAGAGGTGCAAGTGTAGTGGAGACTCCATCCCTGC 29774
Qy 922 ---------- 922
Db 29775 AGGTCCTGTTTCAGGAAATGATTGTACTTTGCCTGAGTAATACAGCCTAGGAGCTACTTT 29834
Qy 922 ---------- 922
Db 29835 CTGATAGGGTTTTTTAAATACTTACAAAGAATTATTTATCTTTAATCATGTGGTTTTGTA 29894
Qy 922 ---------- 922
Db 29895 TGTGTGTGCTTGCACATGCAGTGCTTGTGAGAGAGAGTATGTGTGAGAGCATGCATGTAT 29954
Qy 922 ---------- 922
Db 29955 GAGAGTGTGAGAATATATGTGAGAGAGTGTGAGTGCATGTGTGCGTGTGTGCATCTGTGT 30014
Qy 922 ---------- 922
Db 30015 GTACAGGTGTGTGTACATGCATGTGTGTATAAGAGTATGTGAGAGTGTGGGTGTGTGTGT 30074
Qy 922 ---------- 922
Db 30075 GAGAGTATGTGAGAATATATGTATGAGTGTGTGTGAGTATGAGTGTATGTGCGTGCCTGC 30134
Qy 922 ---------- 922
Db 30135 ATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTAGAAGTGGCCTTGG 30194
Qy 922 ---------- 922
Db 30195 AAAACAGAGTTGTCAGATCTCTTAGAGATATAGTTGCAGTTGGTTGTGAGCCATCTCATA 30254
Qy 922 ---------- 922
Db 30255 TGAGCGCTGGAAGTTGAAATTGGGTTCTCTGGAATCCTCTGGGTTCCTTGTTGAAGCCTG 30314
Qy 922 ---------- 922
Db 30315 AATATTTTGATAAATATTTATGTCATTATCCCTCAAAATTGTAAATGTAGAATTTAACAA 30374
Qy 922 ---------- 922
Db 30375 ACTCAGGTCTTGAGTCATCTTTGTCCCAAGGTTTGTTTGTTTGGTTTTTTGTTCCCCCAC 30434
Qy 922 ---------- 922
Db 30435 CTTTTCTTCAGTGCTTTTAAAAAAGAGAGTCCATTTTTTCCTAAATGTTTAAATACAGTT 30494
Qy 922 ---------- 922
Db 30495 GAGGAATAGAACATCTGACTCCAATTTCCTGGGTTTCCCTCCATGTAGTGTAGTGCTGAC 30554
Qy 922 ---------- 922
Db 30555 CTGATTTCAGTGTGCATTGAAAACTTTGATCACTTGGAAGGCAGCTATGCTCACCACTAT 30614
Qy 922 ---------- 922
Db 30615 ACTACCAATGTCTGCAATCCTATAGGAGAAACAACAATATGAACTAACTAGTACCCCCCA 30674
Qy 922 ---------- 922
Db 30675 GAGCTGTGTCTCTAGTTGCATATGTAGCAGAGGATGGCCTAGTCAGCCATCATTGGGAGG 30734
Qy 922 ---------- 922
Db 30735 AGAGGCCCTTGGTATTGCGAAGATCATATGCCCCAGTACAGGGGAATGCCAGGACCAGGA 30794
Qy 922 ---------- 922


Db 30795 AGCAAGAGTGGGTGGGTTGGGGAGCAGTGCGGGGGGGGGGGGGTATAGGGGGTTTTGGGGA 30854
Qy 922 ---------- 922
Db 30855 TAGCATTTGAAATGTAAATGAAGAAAATAACTAATAAAAATTGCCTTAAAAAAAAACAAA 30914
Qy 923 ----------Asn 923
Db 30915 AAAGAAAAGTTTTTGATCTTAGCTGACCAGTGTCTCTTTGGGTCTTAATTTCCAGCAAAC 30974
Qy 924 HisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThr 943
Db 30975 CACAGTGCCAAGCAGTGCCGGACACCATGTGCCCTGCGGACAGCGTGTGGCGAGTGCACT 31034
Qy 944 SerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAla 963
Db 31035 AGCAGCAGCTCGGAGTGCATGTGGTGCAGTAACATGAAGCAGTGTGTGGACTCCAATGCC 31094
Qy 964 TyrValAlaSerProProPheGlyGlnCysMetGluTrpTyrThrMetSerThrCysPro 983
Db 31095 TACGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCAGCTGCCCA 31154
Qy 983 ---------- 983
Db 31155 CGTAAGTGGAAGGAGCTTTTGAACATTTGCAGGCAAGTTGGGCTTGACTTTCTGCTCAAG 31214
Qy 983 ---------- 983
Db 31215 TCCATGCAGAAGCTGGTCGGGCCGGCCCTTCCAGATTAACATGTATGTATAGAATGCAGC 31274
Qy 983 ---------- 983
Db 31275 ACAGTGTTCCATGCAGTAAATCAGTTACATCAAGGAGAAGGCACAGGGTACAGAAATACC 31334
Qy 983 ---------- 983
Db 31335 TTTTCTTCTTCAGGGTAATATTATAATTCAATCTGTATAATGTTTCTACATCTTAATCTA 31394
Qy 983 ---------- 983
Db 31395 CCAGTATGTAAAGTGCTTTCTAGTAGAGGCCTCCCCAGCTCCCTTTTTCATCCAACATCC 31454
Qy 983 ---------- 983
Db 31455 TGATATTAAAAGGTTGGAAAAGTCCCTGTTATATATTATGTAAAATGTGGGGCCCTTTAA 31514
Qy 983 ---------- 983
Db 31515 ATTATTTCAGTTCAATAATCACTATAGGGTACTATTTTTAATTCATGGAAGTTAAATCAT 31574
Qy 983 ---------- 983
Db 31575 CTGTTAAAAGAAAAGGTAATAACAGTAAATTCAAATCTTGTGATAGTGAATTACAAGTTG 31634
Qy 984 ----------ProGluAsnCysSerGlyTyrCysThrCysSerHi 995
Db 31635 GATTGTTTTGCCTTGTTTTTTAATAGCTGAAAATTGCTCTGGCTACTGTACCTGCAGCCA 31694
Qy 995 sCysLeuGluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLy 1015
Db 31695 TTGCTTGGAGCAGCCAGGCTGTGGTTGGTGTACTGATCCTAGCAATACTGGGAAAGGAAA 31754
Qy 1015 sCysIleGluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAs 1035
Db 31755 ATGTATTGAGGGCAGCTATAAAGGACCTGTGAAGATGCCGTCACAGGCCTCTGCAGGAAA 31814
Qy 1035 nPheTyrProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTr 1055
Db 31815 TGTGTATCCACAGCCCCTTCTGAACTCCAGCATGTGTCTAGAGGACAGCAGATACAACTG 31874
Qy 1055 pSerPheIleHisCysPro 1061

```
Db      31875 GTCTTTCATTCACTGTCCA 31893

RESULT 14
US-09-833-381-1916
; Sequence 1916, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
;  APPLICANT: Robison, Keith E.
;  TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
;  FILE REFERENCE: 5800-119
;  CURRENT APPLICATION NUMBER: US/09/833,381
;  CURRENT FILING DATE:  2001-04-11
;  PRIOR APPLICATION NUMBER: 09/516,448
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 2050
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1916
;   LENGTH: 415
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(415)
;   OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1916

Alignment Scores:
Pred. No.:              1.35e-47      Length:          415
Score:                  581.50        Matches:         105
Percent Similarity:     84.56%        Conservative:    10
Best Local Similarity:  77.21%        Mismatches:      20
Query Match:            7.35%         Indels:          1
DB:                     9             Gaps:            1

US-09-787-097-12 (1-1429) x US-09-833-381-1916 (1-415)

Qy        141 SerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrp 160
              :::||||||::::::|||||||||||  ||||||||||||||||||||||||||||||
Db          4 GCGTCCGGATATTTAACAGATGGCCCAATTAACTATAAATATAAAACTAAATGTACTTGG 63

Qy        161 LeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGlu 180
              ||||||||||||   ||||||   ::::::||||||||||||||||||||||||||||||
Db         64 CTCATTGAAGGCTATCCAAATGCAGTGTTAAGATTAAGATTCAATCATTTTGCTACAGAA 123

Qy        181 CysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAla 200
              ||||||||||||||:::||||||||||||||||||||||||||||||:::|||
Db        124 TGTAGCTGGGATCATATGTATGTTTATGATGGAGATTCAATATATGCACCTTTAATAGCT 183

Qy        201 AlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGluValVal 220
              ||||||||||||||||||||       ||||||||||||||||||||||||||||||
Db        184 GTACTTAGTGGTTTGATAGTCCCTGAAATAAGGGGCAATGAAACTGTGCCTGAAGTTGTT 243

Qy        221 AlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGly 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 ACTACATCTGGCTATGCACTGTTACATTTTTTTAGTGATGCTGCGTATAATCTAACTGGT 303

Qy        241 PheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCys 260
              |||||||||   ||||||   :::   ||||||||||||||||||   |||:::|||
Db        304 TTCAACATTTTCTATTCAATCAATTCTTGTCCTAACAATTGCTCTGGTCATGGGAAGTGT 363

Qy        261 LysIleSerAsnSer---SerGluThrValGluCysGluCysSerGlu 275
                  |||   |||           |||   |||||||||   :::
Db        364 ACAACTAGTGTCTCTGTTCCAAGTCAAGTATATTGTGAATGTGATAAA 411

RESULT 15
US-10-198-846-5869/c
; Sequence 5869, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
;  APPLICANT: Lillie, James
;  APPLICANT:  Xu, Yongyao
;  APPLICANT:  Wang, Youzhen
;  APPLICANT:  Steinmann, Kathleen
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
;  TITLE OF INVENTION:  FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
;  TITLE OF INVENTION:  THERAPY OF BREAST CANCER
;  FILE REFERENCE: MRI-049
;  CURRENT APPLICATION NUMBER: US/10/198,846
;  CURRENT FILING DATE:  2002-07-18
;  PRIOR APPLICATION NUMBER: 60/306,220
;  PRIOR FILING DATE: 2001-07-18
;  NUMBER OF SEQ ID NOS: 14084
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5869
;   LENGTH: 625
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 3, 4, 5, 6, 7, 8, 10, 12, 14, 15, 17, 26, 28, 29
;   OTHER INFORMATION: n = A,T,C or G
US-10-198-846-5869

Alignment Scores:
Pred. No.:              7.9e-43       Length:          625
Score:                  536.00        Matches:         92
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            6.77%         Indels:          0
DB:                     14            Gaps:            0

US-09-787-097-12 (1-1429) x US-10-198-846-5869 (1-625)

Qy        901 ThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArg 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 ACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGG 279

Qy        921 ProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGly 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 CCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGA 219

Qy        941 AspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAsp 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 GATTGCACCAGCGGCAGCTCTGAGTGTATGTGGTGCAGCAACATGAAGCAGTGTGTGGAC 159

Qy        961 SerAsnAlaTyrValAlaSerProProPheGlyGlnCysMetGluTrpTyrThrMetSer 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 TCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGC 99

Qy        981 ThrCysProProGluAsnCysSerGlyTyrCysThr 992
              ||||||||||||||||||||||||||||||
Db         98 ACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACC 63


Search completed: March  1, 2004, 07:35:28
Job time : 1542 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 1, 2004, 00:58:38 ; Search time 235 Seconds
(without alignments)
3374.574 Million cell updates/sec

Title: US-09-787-097-12
Perfect score: 7915
Sequence: 1 MVAAAAATEARLRRRTAATA.....KSGAVRNRKQQPPAQPGTCI 1429

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09787097/runat_26022004_062935_23450/app_query.fasta_1.1607
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09787097_@CGN_1_1_118_@runat_26022004_062935_23450 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	7696.5	97.2	8589	3	US-09-245-041-14	Sequence 14, Appl
	2	7303	92.3	8827	3	US-09-245-041-1	Sequence 1, Appli
	3	6870.5	86.8	4072	3	US-09-245-041-16	Sequence 16, Appl
	4	4136.5	52.3	2625	3	US-09-245-041-18	Sequence 18, Appl
	5	3491.5	44.1	6370	3	US-09-245-041-12	Sequence 12, Appl
	6	3435	43.4	2419	3	US-09-245-041-8	Sequence 8, Appli
	7	1226.5	15.5	1051	3	US-09-245-041-10	Sequence 10, Appl
	8	766	9.7	90050	3	US-09-245-041-5	Sequence 5, Appli
	9	581.5	7.3	415	4	US-09-833-381-1916	Sequence 1916, Ap
c	10	454	5.7	17056	3	US-09-245-041-3	Sequence 3, Appli
c	11	446	5.6	246	4	US-09-702-705-966	Sequence 966, App
c	12	446	5.6	246	4	US-09-736-457-966	Sequence 966, App
c	13	446	5.6	246	4	US-09-614-124B-966	Sequence 966, App
c	14	446	5.6	246	4	US-09-671-325-966	Sequence 966, App
	15	356.5	4.5	9723	1	US-08-083-590A-21	Sequence 21, Appl
	16	356.5	4.5	9723	3	US-08-532-384-21	Sequence 21, Appl
	17	332.5	4.2	8257	4	US-09-484-970B-65	Sequence 65, Appl
	18	326.5	4.1	8091	4	US-09-230-652-1	Sequence 1, Appli
	19	317	4.0	6677	4	US-08-939-366-27	Sequence 27, Appl
	20	316.5	4.0	6677	4	US-09-467-997-6	Sequence 6, Appli
	21	288	3.6	7554	4	US-09-562-702A-29	Sequence 29, Appl
	22	288	3.6	7554	4	US-09-561-818A-25	Sequence 25, Appl
	23	285	3.6	9391	4	US-09-562-702A-11	Sequence 11, Appl
	24	285	3.6	9511	4	US-09-562-702A-9	Sequence 9, Appli
	25	284.5	3.6	5306	4	US-09-562-702A-21	Sequence 21, Appl
	26	284.5	3.6	5306	4	US-09-561-818A-21	Sequence 21, Appl
	27	284.5	3.6	5330	4	US-09-562-702A-25	Sequence 25, Appl
	28	283.5	3.6	11009	4	US-09-845-583A-1	Sequence 1, Appli
	29	281.5	3.6	5184	4	US-09-845-583A-9	Sequence 9, Appli
	30	281.5	3.6	5184	4	US-09-561-709B-4	Sequence 4, Appli
	31	281	3.6	4208	3	US-09-214-278-6	Sequence 6, Appli
	32	281	3.6	4208	4	US-09-855-722-6	Sequence 6, Appli
	33	280	3.5	6464	2	US-08-400-159-5	Sequence 5, Appli
	34	280	3.5	6464	3	US-08-611-729A-5	Sequence 5, Appli
c	35	278	3.5	226	3	US-09-245-041-6	Sequence 6, Appli
	36	278	3.5	5874	4	US-09-561-709B-2	Sequence 2, Appli
	37	277.5	3.5	4948	4	US-09-562-702A-23	Sequence 23, Appl
	38	277.5	3.5	4948	4	US-09-561-818A-23	Sequence 23, Appl
	39	277.5	3.5	4972	4	US-09-562-702A-27	Sequence 27, Appl
	40	277	3.5	4208	4	US-09-068-740A-10	Sequence 10, Appl
	41	277	3.5	5590	3	US-08-882-046-1	Sequence 1, Appli
	42	275	3.5	6942	2	US-08-460-309-3	Sequence 3, Appli
	43	275	3.5	6942	2	US-08-125-077-3	Sequence 3, Appli
	44	275	3.5	9534	4	US-09-562-702A-5	Sequence 5, Appli
	45	275	3.5	9534	4	US-09-561-709B-8	Sequence 8, Appli

ALIGNMENTS

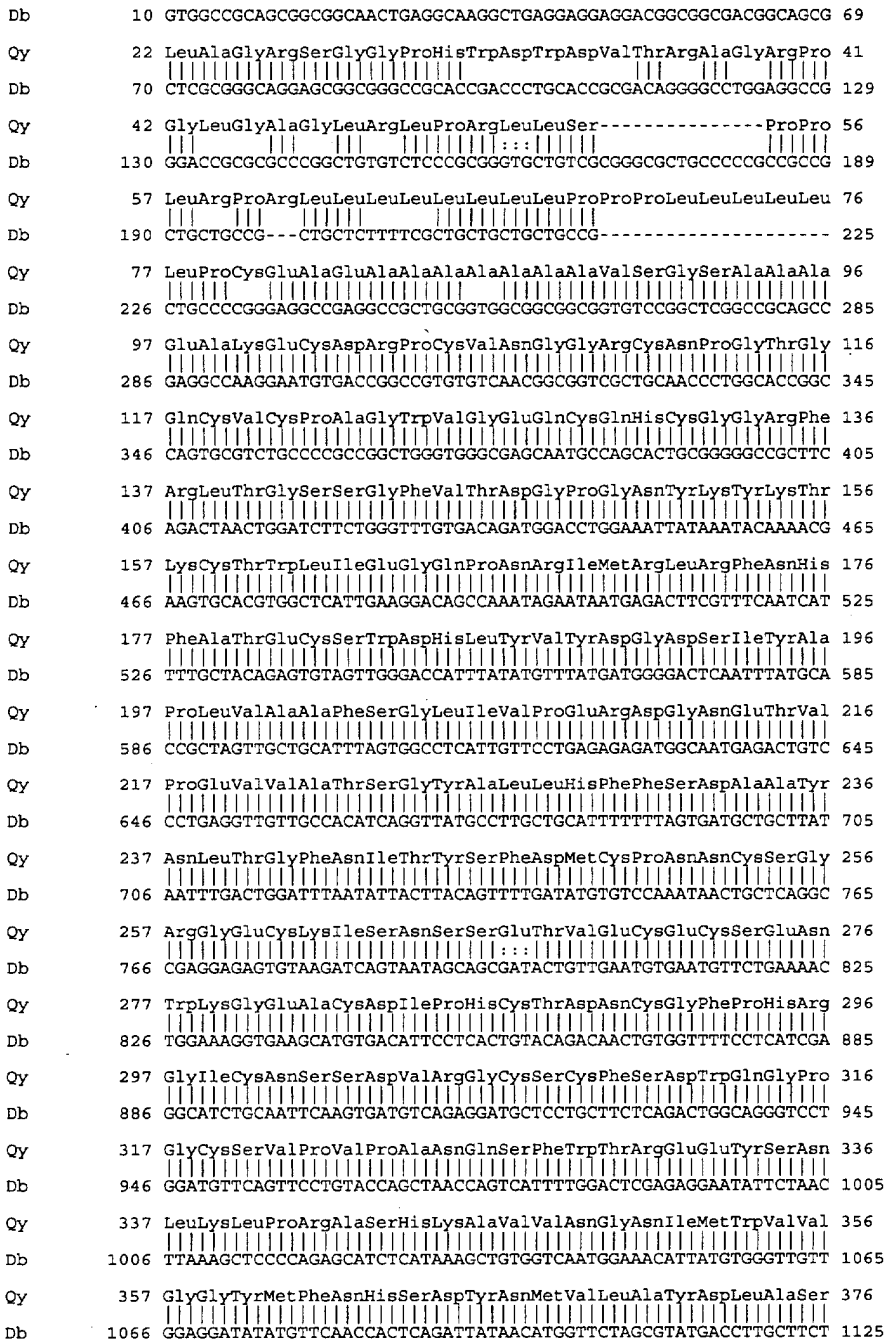
RESULT 1
US-09-245-041-14
; Sequence 14, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 8589
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-245-041-14

Alignment Scores:

Pred. No.:	0	Length:	8589
Score:	7696.50	Matches:	1401
Percent Similarity:	97.91%	Conservative:	2
Best Local Similarity:	97.77%	Mismatches:	17
Query Match:	97.24%	Indels:	13
DB:	3	Gaps:	3

US-09-787-097-12 (1-1429) x US-09-245-041-14 (1-8589)

Qy 2 ValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAlaAla 21


Db 10 GTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCG 69
Qy 22 LeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArgPro 41
Db 70 CTCGCGGGCAGGAGCGGCGGGCCGCACCGACCCTGCACCGCGACAGGGGCCTGGAGGCCG 129
Qy 42 GlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSer----------ProPro 56
Db 130 GGACCGCGCGCCCGGCTGTGTCTCCCGCGGGTGCTGTCGCGGGCGCTGCCCCCGCCGCCG 189
Qy 57 LeuArgProArgLeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeu 76
Db 190 CTGCTGCCG---CTGCTCTTTTCGCTGCTGCTGCTGCCG---------- 225
Qy 77 LeuProCysGluAlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAla 96
Db 226 CTGCCCCGGGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCC 285
Qy 97 GluAlaLysGluCysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGly 116
Db 286 GAGGCCAAGGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGC 345
Qy 117 GlnCysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPhe 136
Db 346 CAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTC 405
Qy 137 ArgLeuThrGlySerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThr 156
Db 406 AGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACG 465
Qy 157 LysCysThrTrpLeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHis 176
Db 466 AAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCAT 525
Qy 177 PheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAla 196
Db 526 TTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCA 585
Qy 197 ProLeuValAlaAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrVal 216
Db 586 CCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTC 645
Qy 217 ProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyr 236
Db 646 CCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTAT 705
Qy 237 AsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGly 256
Db 706 AATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGC 765
Qy 257 ArgGlyGluCysLysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsn 276
Db 766 CGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAAC 825
Qy 277 TrpLysGlyGluAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArg 296
Db 826 TGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGA 885
Qy 297 GlyIleCysAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyPro 316
Db 886 GGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCT 945
Qy 317 GlyCysSerValProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsn 336
Db 946 GGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAAC 1005
Qy 337 LeuLysLeuProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValVal 356
Db 1006 TTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTT 1065
Qy 357 GlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSer 376
Db 1066 GGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCT 1125
Qy 377 ArgGluTrpLeuProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSer 396
Db 1126 AGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCT 1185
Qy 397 LeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsn 416
Db 1186 TTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAAT 1245
Qy 417 ValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrPro 436
Db 1246 GTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCT 1305
Qy 437 LysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsn 456
Db 1306 AAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAAT 1365
Qy 457 GlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsn 476
Db 1366 GGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAAT 1425
Qy 477 ValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeu 496
Db 1426 GTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTT 1485
Qy 497 ValGlnGlyGlyTyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrVal 516
Db 1486 GTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTT 1545
Qy 517 HisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArg 536
Db 1546 CATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGA 1605
Qy 537 TyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeu 556
Db 1606 TATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTG 1665
Qy 557 HisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAsp 576
Db 1666 CACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGAC 1725
Qy 577 ThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAla 596
Db 1726 ACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCC 1785
Qy 597 CysAspArgTrpSerValLeuProArgProAspLeuHisHisAspValAsnArgPheGly 616
Db 1786 TGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGC 1845
Qy 617 HisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeu 636
Db 1846 CATTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTC 1905
Qy 637 LeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAla 656
Db 1906 CTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCT 1965
Qy 657 CysLeuAlaAlaGlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIle 676
Db 1966 TGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATC 2025
Qy 677 SerTrpAlaLeuAlaThrAspGluGlnGluGluLysLeuLysSerGluCysPheSerLys 696
Db 2026 TCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAA 2085
Qy 697 ArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsn 716
Db 2086 AGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAAC 2145
Qy 717 ThrAsnAspCysHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGlu 736
Db 2146 ACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAA 2205

```
Qy       737 GlyGlnIleSerIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCys 756
Db      2206 GGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGT 2265

Qy       757 AsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArg 776
Db      2266 AACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGG 2325

Qy       777 AsnGlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGly 796
Db      2326 AATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGA 2385

Qy       797 AsnSerCysLeuLysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCys 816
Db      2386 AACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGT 2445

Qy       817 ArgAsnHisAsnAlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeu 836
Db      2446 AGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTT 2505

Qy       837 LysGlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrp 856
Db      2506 AAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGG 2565

Qy       857 ValGlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThr 876
Db      2566 GTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACA 2625

Qy       877 AsnSerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeu 896
Db      2626 AATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTA 2685

Qy       897 SerGluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySer 916
Db      2686 TCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGT 2745

Qy       917 ValCysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArg 936
Db      2746 GTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGG 2805

Qy       937 ThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLys 956
Db      2806 ACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAG 2865

Qy       957 GlnCysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrp 976
Db      2866 CAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGG 2925

Qy       977 TyrThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCys 996
Db      2926 TATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGC 2985

Qy       997 LeuGluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCys 1016
Db      2986 TTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGC 3045

Qy      1017 IleGluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPhe 1036
Db      3046 ATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTC 3105

Qy      1037 TyrProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSer 1056
Db      3106 TATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCT 3165

Qy      1057 PheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIle 1076
Db      3166 TTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATC 3225

Qy      1077 CysGluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPhe 1096
Db      3226 TGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTC 3285

Qy      1097 TyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSer 1116
Db      3286 TACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCT 3345

Qy      1117 LeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGlu 1136
Db      3346 CTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAG 3405

Qy      1137 CysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyr 1156
Db      3406 TGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTAT 3465

Qy      1157 TyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyr 1176
Db      3466 TATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTAT 3525

Qy      1177 TyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPhe 1196
Db      3526 TACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTC 3585

Qy      1197 IleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGly 1216
Db      3586 ATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGA 3645

Qy      1217 ThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAsp 1236
Db      3646 ACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGAT 3705

Qy      1237 SerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyr 1256
Db      3706 AGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTAT 3765

Qy      1257 ValSerAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPhe 1276
Db      3766 GTCAGTAATTTCACCTGGCCCATCAAAATTCAGATTGCCTTCTCTCAGCACAGCAATTTT 3825

Qy      1277 MetAspLeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuLeuVal 1296
Db      3826 ATGGACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTG 3885

Qy      1297 AlaAlaValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeu 1316
Db      3886 GCTGCTGTGGTTTGGAAGATCAAACAAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTT 3945

Qy      1317 LeuArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGlu 1336
Db      3946 CTTCGAGAGATGCAACAGATGGCCAGCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAA 4005

Qy      1337 ThrAspGluGluProProAspLeuIleGlyGlySerIleLysThrValProLysProIle 1356
Db      4006 ACAGATGAGGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATT 4065

Qy      1357 AlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuPro 1376
Db      4066 GCACTGGAGCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCT 4125

Qy      1377 ArgGlyLeuGlyGlyIleProProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeu 1396
Db      4126 CGAGGCCTGGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTG 4185

Qy      1397 ValAspIleSerGlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAsn 1416
Db      4186 GTGGACATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAAC 4245

Qy      1417 ArgLysGlnGlnProProAlaGlnProGlyThrCysIle 1429
Db      4246 CGGAAGCAGCAGCCCCCTGCACAGCCTGGGACCTGCATC 4284
```

RESULT 2
US-09-245-041-1
; Sequence 1, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.

```
;  APPLICANT: Nagle, D.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
;  TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
;  FILE REFERENCE: 7853-136
;  CURRENT APPLICATION NUMBER: US/09/245,041
;  CURRENT FILING DATE: 1999-02-05
;  EARLIER APPLICATION NUMBER: 60/093,630
;  EARLIER FILING DATE: 1998-07-21
;  EARLIER APPLICATION NUMBER: 60/104,978
;  EARLIER FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 131
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 8827
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-245-041-1

Alignment Scores:
Pred. No.:                0              Length:          8827
Score:                    7303.00        Matches:         1332
Percent Similarity:       90.99%         Conservative:    32
Best Local Similarity:    88.86%         Mismatches:      56
Query Match:              92.27%         Indels:          79
DB:                       3              Gaps:            4

US-09-787-097-12 (1-1429) x US-09-245-041-1 (1-8827)

Qy          2 ValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAlaAla 21
              ||||||||||||||||||||||||||||||||||  |||  |||||||||
Db         83 GTGGCGGCGGCGGCGGCGACTGAGGCGCGGCTGAGGGGGAGCACGAGGACGACAGCAGCG 142

Qy         22 LeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArgPro 41
                 |||||||||   |||     |||          |||   |||   ||||||
Db        143 CCTGCGGGCAGGAAGGGCAGGCAGCACCGACCCTGCACCGCGACAGGGGCCTGGAGGCCG 202

Qy         42 GlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSer----------------ProPro 56
              |||       |||   |||   |||||||||:::||||||                ||||||
Db        203 GGACCGCGCGCCCGGCTGTGTCTCCCGCGGGTGCTGTCGCGGGCGCTGCCCCCGCCGCCG 262

Qy         57 LeuArgProArgLeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeu 76
              |||   |||   ||||||      ||||||||||||||||
Db        263 CTGCTGCCG---CTGCTCTTTTCGCTGCTGCTGCTGCCG---------------------- 298

Qy         77 LeuProCysGluAlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAla 96
              ||||||   ||||||||||||||||||   ||||||||||||||||||||||||||||||
Db        299 CTGCCCCGGGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCC 358

Qy         97 GluAlaLysGluCysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGly 116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 GAGGCCAAGGAATGTGACCGGCCGTGTGTCAACGGCGGCCGCTGCAACCCTGGCACCGGC 418

Qy        117 GlnCysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPhe 136
              ||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db        419 CAGTGCGTCTGCCCCACGGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTC 478

Qy        136 ------------------------------------------------------------ 136

Db        479 AGGACATCTGTCTCACGCCTATAATCACAGCTGTTCGGAAGGTGAGGCTGGAGGAACAGT 538

Qy        136 ------------------------------------------------------------ 136

Db        539 TCGAGGCAAGCTTCGGCTACAGAATAAGTTCAAGAGTAACCTGGGGCAACTTGGGCTTGT 598

Qy        136 ------------------------------------------------------------ 136

Db        599 CTCCAAAACCAAAATGAGCGAAAAGGAGCAAGCTAGAGTCTTTTGGGAAAATTTTAGCTG 658

Qy        137 --------------Arg-LeuThrGlySerSerGlyPheValThrAspGlyProGlyAs 151
                           ||| ||||||||||||||||||||||||||||||||||||||||||
Db        659 ACTAATTTTTCACCGAGAACTAACTGGCTCTTCTGGATTTGTAACAGATGGACCTGGGAA 718

Qy        151 nTyrLysTyrLysThrLysCysThrTrpLeuIleGluGlyGlnProAsnArgIleMetAr 171
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 TTATAAATATAAGACGAAGTGCACATGGCTCATTGAAGGACAGCCAAATAGAATAATGAG 778

Qy        171 gLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGl 191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 ACTTCGCTTCAACCATTTTGCTACAGAATGTAGCTGGGACCATTTATATGTTTATGATGG 838

Qy        191 yAspSerIleTyrAlaProLeuValAlaAlaPheSerGlyLeuIleValProGluArgAs 211
              ||||||||||||||||||||||:::||||||||||||||||||||||||||||||||||||
Db        839 GGACTCAATCTACGCACCTCTGATTGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGA 898

Qy        211 pGlyAsnGluThrValProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePh 231
              |||||||||||||   |||||||||      ||||||||||||||||||||||||||||||
Db        899 TGGCAATGAGACGGCTCCTGAGGTCACTGTCACTTCAGGTTATGCACTGCTGCATTTTTT 958

Qy        231 eSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysPr 251
              ||||||||||||||||||||||||||||||||||||||||||:::||||||||||||||||
Db        959 CAGTGATGCTGCTTATAATCTGACTGGATTTAATATCACTTACAATTTTGACATGTGTCC 1018

Qy        251 oAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSerSerGluThrValGluCy 271
              |||||||||||||||||||||||||||||   ||||||||||||      ||||||||
Db       1019 GAATAATTGCTCAGGCCGAGGAGAGTGTAAGAGCAGTAACAGCAGCAGCGCTGTTGAGTG 1078

Qy        271 sGluCysSerGluAsnTrpLysGlyGluAlaCysAspIleProHisCysThrAspAsnCy 291
              ||||||||||||||||||||||||||:::||||||||||||||||||||||||||||||||
Db       1079 TGAATGTTCTGAAAACTGGAAAGGGGAGTCGTGTGACATTCCTCACTGTACAGACAACTG 1138

Qy        291 sGlyPheProHisArgGlyIleCysAsnSerSerAspValArgGlyCysSerCysPheSe 311
              ||||||||||||||||||||||||||||:::||||||    ||||||||||||||||||||
Db       1139 TGGCTTTCCTCACCGAGGCATCTGTAATGCAAGCGATACCAGAGGGTGCTCCTGCTTTCC 1198

Qy        311 rAspTrpGlnGlyProGlyCysSerValProValProAlaAsnGlnSerPheTrpThrAr 331
                   ||||||||||||||||||||:::||||||||||||||||||||||||||||||||||
Db       1199 TCACTGGCAGGGTCCTGGATGTTCAATTCCTGTGCCAGCTAACCAGTCTTTTTGGACTCG 1258

Qy        331 gGluGluTyrSerAsnLeuLysLeuProArgAlaSerHisLysAlaValValAsnGlyAs 351
              ||||||||||||:::||||||||||||||||||||||||||||||||||||||||||||||
Db       1259 AGAAGAATATTCTGATTTAAAGCTTCCCAGAGCCTCTCATAAAGCTGTGGTCAATGGAAA 1318

Qy        351 nIleMetTrpValValGlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAl 371
              ||||||||||||||||||||||||||||||||||||||||||||||||:::||||||||||
Db       1319 TATAATGTGGGTTGTTGGCGGATATATGTTCAACCATTCAGATTACAGCATGGTTTTAGC 1378

Qy        371 aTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSerValAsnAsnValValVa 391
              ||||||||||   ||||||||||||||||||||||||   |||||||||:::||||||||
Db       1379 GTATGACCTGACTTCTAGGGAATGGCTTCCACTAAACCATTCTGTGAACAGTGTGGTTGT 1438

Qy        391 lArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLysIl 411
              ||||||||||||||||||||||||||||:::||||||||||||||||||||||||||||||
Db       1439 AAGATATGGTCATTCTTTGGCATTACATAAGGATAAAATCTACATGTATGGAGGAAAAAT 1498

Qy        411 eAspProThrGlyAsnValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTr 431
              ||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1499 TGATTCAACAGGGAACGTGACCAATGAGCTGAGAGTATTTCATATTCATAATGAATCATG 1558

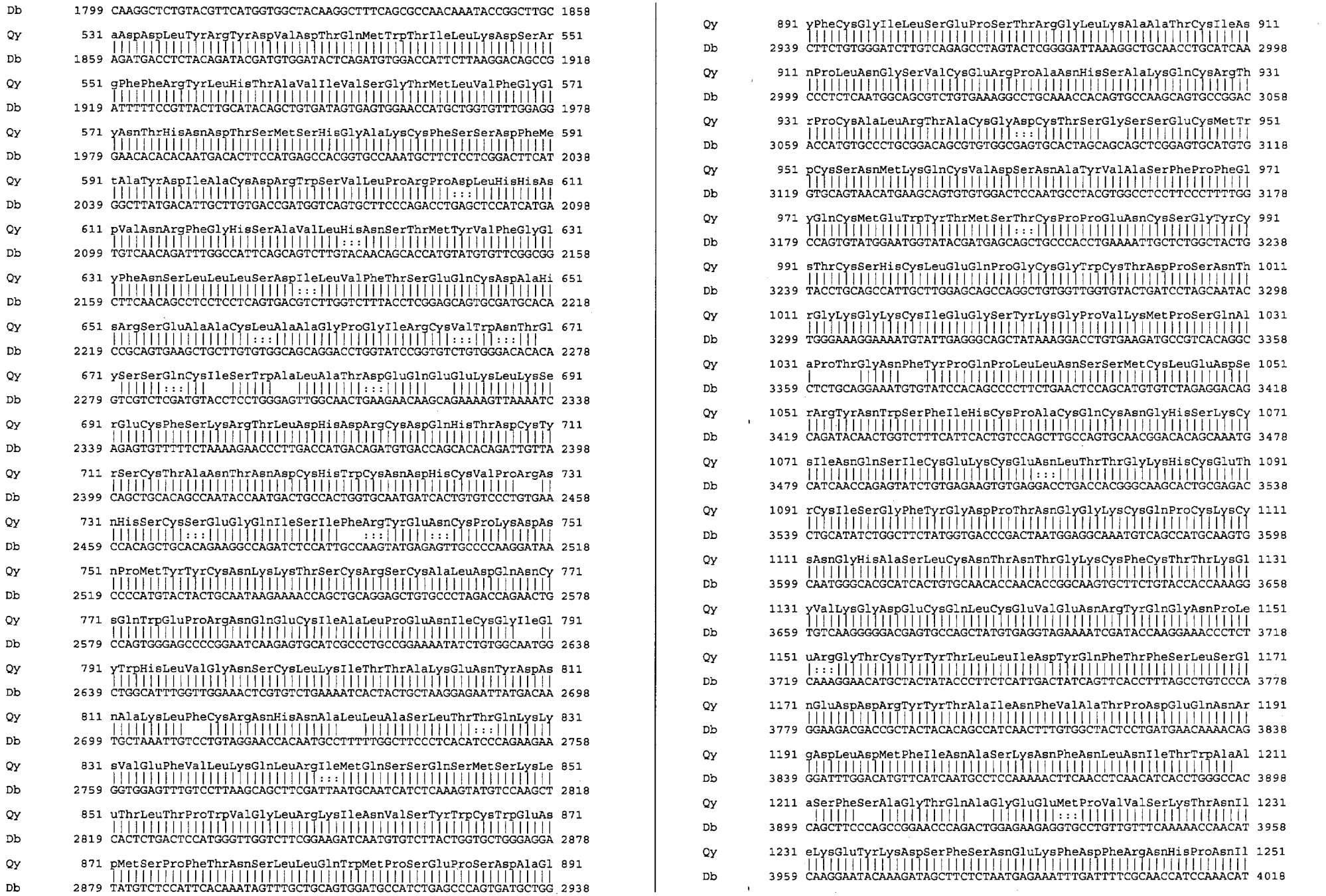
Qy        431 pValLeuLeuThrProLysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIl 451
              ||||||||||||||||||||||||:::||||||||||||||||||||||||||||||||||
Db       1559 GGTATTGTTAACTCCGAAAGCTAAGGATCAGTATGCAGTGGTTGGACACTCAGCACACAT 1618

Qy        451 eValThrLeuLysAsnGlyArgValValMetLeuValIlePheGlyHisCysProLeuTy 471
              ||||||||||   :::||||||||||||||||||||||||||||||||||||||||||||||
Db       1619 TGTTACACTGGCATCTGGCCGTGTGGTCATGTTGGTCATCTTCGGTCATTGCCCACTCTA 1678

Qy        471 rGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHi 491
              |||||||||||||   ||||||||||||||||||:::||||||||||||||||||||||||
Db       1679 TGGATATATAAGCGTTGTGCAGGAATATGACTTGGAAAAGAACACATGGAGTATATTACA 1738

Qy        491 sThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisSerSerValTyrAspHisArgTh 511
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1739 TACTCAGGGTGCTCTTGTGCAAGGGGGTTATGGCCACAGTAGTGTTTATGATGACAGGAC 1798

Qy        511 rArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAl 531
              |:::||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```


Db 1799 CAAGGCTCTGTACGTTCATGGTGGCTACAAGGCTTTCAGCGCCAACAAATACCGGCTTGC 1858
Qy 531 aAspAspLeuTyrArgTyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerAr 551
Db 1859 AGATGACCTCTACAGATACGATGTGGATACTCAGATGTGGACCATTCTTAAGGACAGCCG 1918
Qy 551 gPhePheArgTyrLeuHisThrAlaValIleValSerGlyThrMetLeuValPheGlyGl 571
Db 1919 ATTTTTCCGTTACTTGCATACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGAGG 1978
Qy 571 yAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMe 591
Db 1979 GAACACACACAATGACACTTCCATGAGCCACGGTGCCAAATGCTTCTCCTCGGACTTCAT 2038
Qy 591 tAlaTyrAspIleAlaCysAspArgTrpSerValLeuProArgProAspLeuHisHisAs 611
Db 2039 GGCTTATGACATTGCTTGTGACCGATGGTCAGTGCTTCCCAGACCTGAGCTCCATCATGA 2098
Qy 611 pValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGl 631
Db 2099 TGTCAACAGATTTGGCCATTCAGCAGTCTTGTACAACAGCACCATGTATGTGTTCGGCGG 2158
Qy 631 yPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHi 651
Db 2159 CTTCAACAGCCTCCTCCTCAGTGACGTCTTGGTCTTTACCTCGGAGCAGTGCGATGCACA 2218
Qy 651 sArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgCysValTrpAsnThrGl 671
Db 2219 CCGCAGTGAAGCTGCTTGTGTGGCAGCAGGACCTGGTATCCGGTGTCTGTGGGACACACA 2278
Qy 671 ySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnGluGluLysLeuLysSe 691
Db 2279 GTCGTCTCGATGTACCTCCTGGGAGTTGGCAACTGAAGAACAAGCAGAAAAGTTAAAATC 2338
Qy 691 rGluCysPheSerLysArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTy 711
Db 2339 AGAGTGTTTTTCTAAAAGAACCCTTGACCATGACAGATGTGACCAGCACACAGATTGTTA 2398
Qy 711 rSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAspHisCysValProArgAs 731
Db 2399 CAGCTGCACAGCCAATACCAATGACTGCCACTGGTGCAATGATCACTGTGTCCCTGTGAA 2458
Qy 731 nHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGluAsnCysProLysAspAs 751
Db 2459 CCACAGCTGCACAGAAGGCCAGATCTCCATTGCCAAGTATGAGAGTTGCCCCAAGGATAA 2518
Qy 751 nProMetTyrTyrCysAsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCy 771
Db 2519 CCCCATGTACTACTGCAATAAGAAAACCAGCTGCAGGAGCTGTGCCCTAGACCAGAACTG 2578
Qy 771 sGlnTrpGluProArgAsnGlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGl 791
Db 2579 CCAGTGGGAGCCCCGGAATCAAGAGTGCATCGCCCTGCCGGAAAATATCTGTGGCAATGG 2638
Qy 791 yTrpHisLeuValGlyAsnSerCysLeuLysIleThrThrAlaLysGluAsnTyrAspAs 811
Db 2639 CTGGCATTTGGTTGGAAACTCGTGTCTGAAAATCACTACTGCTAAGGAGAATTATGACAA 2698
Qy 811 nAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSerLeuThrThrGlnLysLy 831
Db 2699 TGCTAAATTGTCCTGTAGGAACCACAATGCCTTTTTGGCTTCCCTCACATCCCAGAAGAA 2758
Qy 831 sValGluPheValLeuLysGlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLe 851
Db 2759 GGTGGAGTTTGTCCTTAAGCAGCTTCGATTAATGCAATCATCTCAAAGTATGTCCAAGCT 2818
Qy 851 uThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAs 871
Db 2819 CACTCTGACTCCATGGGTTGGTCTTCGGAAGATCAATGTGTCTTACTGGTGCTGGGAGGA 2878
Qy 871 pMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSerAspAlaGl 891
Db 2879 TATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGTGATGCTGG 2938
Qy 891 yPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThrCysIleAs 911
Db 2939 CTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGATTAAAGGCTGCAACCTGCATCAA 2998
Qy 911 nProLeuAsnGlySerValCysGluArgProAlaAsnHisSerAlaLysGlnCysArgTh 931
Db 2999 CCCTCTCAATGGCAGCGTCTGTGAAAGGCCTGCAAACCACAGTGCCAAGCAGTGCCGGAC 3058
Qy 931 rProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTr 951
Db 3059 ACCATGTGCCCTGCGGACAGCGTGTGGCGAGTGCACTAGCAGCAGCTCGGAGTGCATGTG 3118
Qy 951 pCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrValAlaSerProPheGl 971
Db 3119 GTGCAGTAACATGAAGCAGTGTGTGGACTCCAATGCCTACGTGGCCTCCTTCCCTTTTGG 3178
Qy 971 yGlnCysMetGluTrpTyrThrMetSerThrCysProProGluAsnCysSerGlyTyrCy 991
Db 3179 CCAGTGTATGGAATGGTATACGATGAGCAGCTGCCCACCTGAAAATTGCTCTGGCTACTG 3238
Qy 991 sThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCysThrAspProSerAsnTh 1011
Db 3239 TACCTGCAGCCATTGCTTGGAGCAGCCAGGCTGTGGTTGGTGTACTGATCCTAGCAATAC 3298
Qy 1011 rGlyLysGlyLysCysIleGluGlySerTyrLysGlyProValLysMetProSerGlnAl 1031
Db 3299 TGGGAAAGGAAAATGTATTGAGGGCAGCTATAAAGGACCTGTGAAGATGCCGTCACAGGC 3358
Qy 1031 aProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSe 1051
Db 3359 CTCTGCAGGAAATGTGTATCCACAGCCCCTTCTGAACTCCAGCATGTGTCTAGAGGACAG 3418
Qy 1051 rArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCy 1071
Db 3419 CAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAGTGCAACGGACACAGCAAATG 3478
Qy 1071 sIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrThrGlyLysHisCysGluTh 1091
Db 3479 CATCAACCAGAGTATCTGTGAGAAGTGTGAGGACCTGACCACGGGCAAGCACTGCGAGAC 3538
Qy 1091 rCysIleSerGlyPheTyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLysCy 1111
Db 3539 CTGCATATCTGGCTTCTATGGTGACCCGACTAATGGAGGCAAATGTCAGCCATGCAAGTG 3598
Qy 1111 sAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGl 1131
Db 3599 CAATGGGCACGCATCACTGTGCAACACCAACACCGGCAAGTGCTTCTGTACCACCAAAGG 3658
Qy 1131 yValLysGlyAspGluCysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLe 1151
Db 3659 TGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCT 3718
Qy 1151 uArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGl 1171
Db 3719 CAAAGGAACATGCTACTATACCCTTCTCATTGACTATCAGTTCACCTTTAGCCTGTCCCA 3778
Qy 1171 nGluAspAspArgTyrTyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnAr 1191
Db 3779 GGAAGACGACCGCTACTACACAGCCATCAACTTTGTGGCTACTCCTGATGAACAAAACAG 3838
Qy 1191 gAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAl 1211
Db 3839 GGATTTGGACATGTTCATCAATGCCTCCAAAAACTTCAACCTCAACATCACCTGGGCCAC 3898
Qy 1211 aSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIl 1231
Db 3899 CAGCTTCCCAGCCGGAACCCAGACTGGAGAAGAGGTGCCTGTTGTTTCAAAAACCAACAT 3958
Qy 1231 eLysGluTyrLysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIl 1251
Db 3959 CAAGGAATACAAAGATAGCTTCTCTAATGAGAAATTTGATTTTCGCAACCATCCAAACAT 4018

```
Qy       1251 eThrPhePheValTyrValSerAsnPheThrTrpProIleLysIleGlnIleAlaPheSe 1271
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4019 CACTTTCTTTGTTTATGTCAGTAATTTCACTTGGCCCATCAAAATTCAGATTGCCTTCTC 4078

Qy       1271 rGlnHisSerAsnPheMetAspLeuValGlnPhePheValThrPhePheSerCysPheLe 1291
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4079 CCAGCACAGCAACTTCATGGACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCT 4138

Qy       1291 uSerLeuLeuLeuValAlaAlaValValTrpLysIleLysGlnSerCysTrpAlaSerAr 1311
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4139 CTCGCTGCTTCTGGTGGCTGCAGTGGTCTGGAAGATCAAGCAGAGCTGTTGGGCATCCAG 4198

Qy       1311 gArgArgGluGlnLeuLeuArgGluMetGlnGlnMetAlaSerArgProPheAlaSerVa 1331
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4199 GCGGAGAGAGCAACTTCTTCGGGAGATGCAACAGATGGCCAGCCGCCCCTTTGCTTCTGT 4258

Qy       1331 lAsnValAlaLeuGluThrAspGluGluProProAspLeuIleGlyGlySerIleLysTh 1351
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4259 AAACGTTGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGAC 4318

Qy       1351 rValProLysProIleAlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerVa 1371
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4319 TGTTCCCAAACCCATTGCACTGGAGCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGT 4378

Qy       1371 lPheValArgLeuProArgGlyLeuGlyGlyIleProProProGlyGlnSerGlyLeuAl 1391
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4379 GTTTGTGAGGCTCCCTCGAGGCCTGGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGC 4438

Qy       1391 aValAlaSerAlaLeuValAspIleSerGlnGlnMetProIleValTyrLysGluLysSe 1411
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4439 TGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTC 4498

Qy       1411 rGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnProGlyThrCysIle 1429
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4499 AGGAGCCGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGCCTGGGACCTGCATC 4553

RESULT 3
US-09-245-041-16
; Sequence 16, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT: Nagle, D.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
;  TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
;  FILE REFERENCE: 7853-136
;  CURRENT APPLICATION NUMBER: US/09/245,041
;  CURRENT FILING DATE: 1999-02-05
;  EARLIER APPLICATION NUMBER: 60/093,630
;  EARLIER FILING DATE: 1998-07-21
;  EARLIER APPLICATION NUMBER: 60/104,978
;  EARLIER FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 131
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 16
;   LENGTH: 4072
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-245-041-16

Alignment Scores:
Pred. No.:              0           Length:          4072
Score:                  6870.50     Matches:         1240
Percent Similarity:     97.41%      Conservative:    3
Best Local Similarity:  97.18%      Mismatches:      20
Query Match:            86.80%      Indels:          13
DB:                     3           Gaps:            3

US-09-787-097-12 (1-1429) x US-09-245-041-16 (1-4072)

Qy          2 ValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAlaAla 21
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         10 GTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCG 69

Qy         22 LeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArgPro 41
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 CTCGCGGGCAGGAGCGGCGGGCCGCACCGACCCTGCACCGCGACAGGGGCCTGGAGGCCG 129

Qy         42 GlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSer----------------ProPro 56
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 GGACCGCGCGCCCGGCTGTGTCTCCCGCGGGTGCTGTCGCGGGCGCTGCCCCCGCCGCCG 189

Qy         57 LeuArgProArgLeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeu 76
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 CTGCTGCCG---CTGCTCTTTTCGCTGCTGCTGCTGCCG---------------------- 225

Qy         77 LeuProCysGluAlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAla 96
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        226 CTGCCCCGGGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCC 285

Qy         97 GluAlaLysGluCysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGly 116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 GAGGCCAAGGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGC 345

Qy        117 GlnCysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPhe 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 CAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTC 405

Qy        137 ArgLeuThrGlySerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThr 156
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        406 AGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACG 465

Qy        157 LysCysThrTrpLeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHis 176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        466 AAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCAT 525

Qy        177 PheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAla 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        526 TTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCA 585

Qy        197 ProLeuValAlaAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrVal 216
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        586 CCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTC 645

Qy        217 ProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyr 236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        646 CCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTAT 705

Qy        237 AsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGly 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        706 AATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGC 765

Qy        257 ArgGlyGluCysLysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsn 276
              ||||||||||||||||||||||||||:::||||||||||||||||||||||||||||||||
Db        766 CGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAAC 825

Qy        277 TrpLysGlyGluAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArg 296
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        826 TGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGA 885

Qy        297 GlyIleCysAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyPro 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        886 GGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCT 945

Qy        317 GlyCysSerValProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsn 336
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        946 GGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAAC 1005

Qy        337 LeuLysLeuProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValVal 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1006 TTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTT 1065

Qy        357 GlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSer 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1066 GGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCT 1125
```

```
Qy      377 ArgGluTrpLeuProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSer 396
Db     1126 AGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCT 1185

Qy      397 LeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsn 416
Db     1186 TTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAAT 1245

Qy      417 ValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrPro 436
Db     1246 GTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCT 1305

Qy      437 LysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsn 456
Db     1306 AAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAAT 1365

Qy      457 GlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsn 476
Db     1366 GGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAAT 1425

Qy      477 ValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeu 496
Db     1426 GTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTT 1485

Qy      497 ValGlnGlyGlyTyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrVal 516
Db     1486 GTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTT 1545

Qy      517 HisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArg 536
Db     1546 CATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGA 1605

Qy      537 TyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeu 556
Db     1606 TATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTG 1665

Qy      557 HisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAsp 576
Db     1666 CACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGAC 1725

Qy      577 ThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAla 596
Db     1726 ACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCC 1785

Qy      597 CysAspArgTrpSerValLeuProArgProAspLeuHisHisAspValAsnArgPheGly 616
Db     1786 TGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGC 1845

Qy      617 HisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeu 636
Db     1846 CATTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTC 1905

Qy      637 LeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAla 656
Db     1906 CTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCT 1965

Qy      657 CysLeuAlaAlaGlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIle 676
Db     1966 TGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATC 2025

Qy      677 SerTrpAlaLeuAlaThrAspGluGlnGluGluLysLeuLysSerGluCysPheSerLys 696
Db     2026 TCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAA 2085

Qy      697 ArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsn 716
Db     2086 AGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAAC 2145

Qy      717 ThrAsnAspCysHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGlu 736
Db     2146 ACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAA 2205

Qy      737 GlyGlnIleSerIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCys 756

Db     2206 GGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGT 2265

Qy      757 AsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArg 776
Db     2266 AACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGG 2325

Qy      777 AsnGlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGly 796
Db     2326 AATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGA 2385

Qy      797 AsnSerCysLeuLysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCys 816
Db     2386 AACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGT 2445

Qy      817 ArgAsnHisAsnAlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeu 836
Db     2446 AGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTT 2505

Qy      837 LysGlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrp 856
Db     2506 AAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGG 2565

Qy      857 ValGlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThr 876
Db     2566 GTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACA 2625

Qy      877 AsnSerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeu 896
Db     2626 AATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTA 2685

Qy      897 SerGluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySer 916
Db     2686 TCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGT 2745

Qy      917 ValCysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArg 936
Db     2746 GTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGG 2805

Qy      937 ThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLys 956
Db     2806 ACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAG 2865

Qy      957 GlnCysValAspSerAsnAlaTyrValAlaSerProProPheGlyGlnCysMetGluTrp 976
Db     2866 CAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGG 2925

Qy      977 TyrThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCys 996
Db     2926 TATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGC 2985

Qy      997 LeuGluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCys 1016
Db     2986 TTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGC 3045

Qy     1017 IleGluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPhe 1036
Db     3046 ATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTC 3105

Qy     1037 TyrProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSer 1056
Db     3106 TATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCT 3165

Qy     1057 PheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIle 1076
Db     3166 TTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATC 3225

Qy     1077 CysGluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPhe 1096
Db     3226 TGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTC 3285

Qy     1097 TyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSer 1116
```

```
Db       3286 TACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCT 3345

Qy       1117 LeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGlu 1136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3346 CTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAG 3405

Qy       1137 CysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyr 1156
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3406 TGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTAT 3465

Qy       1157 TyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyr 1176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3466 TATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTAT 3525

Qy       1177 TyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPhe 1196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3526 TACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTC 3585

Qy       1197 IleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGly 1216
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3586 ATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGA 3645

Qy       1217 ThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAsp 1236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3646 ACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGAT 3705

Qy       1237 SerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyr 1256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3706 AGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTAT 3765

Qy       1257 ValSerAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGln 1272
              ||||||||||||||||||||||||||||||||:::   |||
Db       3766 GTCAGTAATTTCACCTGGCCCATCAAAATTCAGGTGCAAACTGAACAA 3813

RESULT 4
US-09-245-041-18
; Sequence 18, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT: Nagle, D.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
;  TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
;  FILE REFERENCE: 7853-136
;  CURRENT APPLICATION NUMBER: US/09/245,041
;  CURRENT FILING DATE: 1999-02-05
;  EARLIER APPLICATION NUMBER: 60/093,630
;  EARLIER FILING DATE: 1998-07-21
;  EARLIER APPLICATION NUMBER: 60/104,978
;  EARLIER FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 131
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 18
;   LENGTH: 2625
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-245-041-18

Alignment Scores:
Pred. No.:              0              Length:          2625
Score:                  4136.50        Matches:         759
Percent Similarity:     95.48%         Conservative:    2
Best Local Similarity:  95.23%         Mismatches:      22
Query Match:            52.26%         Indels:          14
DB:                     3              Gaps:            3

US-09-787-097-12 (1-1429) x US-09-245-041-18 (1-2625)

Qy          2 ValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAlaAla 21
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         10 GTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCAGCG 69

Qy         22 LeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArgPro 41
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 CTCGCGGGCAGGAGCGGCGGGCCGCACCGACCCTGCACCGCGACAGGGGCCTGGAGGCCG 129

Qy         42 GlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSer---------------ProPro 56
              |||   |||  |||  |||||||||:::||||||                  ||||||
Db        130 GGACCGCGCGCCCGGCTGTGTCTCCCGCGGGTGCTGTCGCGGGCGCTGCCCCCGCCGCCG 189

Qy         57 LeuArgProArgLeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeu 76
              |||   |||  ||||||        ||||||||||||||||
Db        190 CTGCTGCCG---CTGCTCTTTTCGCTGCTGCTGCTGCCG---------------------- 225

Qy         77 LeuProCysGluAlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAla 96
              ||||||  ||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db        226 CTGCCCCGGGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCC 285

Qy         97 GluAlaLysGluCysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGly 116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 GAGGCCAAGGAATGTGACCGGCCGTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGC 345

Qy        117 GlnCysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPhe 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 CAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTC 405

Qy        137 ArgLeuThrGlySerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThr 156
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        406 AGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACG 465

Qy        157 LysCysThrTrpLeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHis 176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        466 AAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCAT 525

Qy        177 PheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAla 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        526 TTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCA 585

Qy        197 ProLeuValAlaAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrVal 216
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        586 CCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTC 645

Qy        217 ProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyr 236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        646 CCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTAT 705

Qy        237 AsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGly 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        706 AATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGC 765

Qy        257 ArgGlyGluCysLysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsn 276
              ||||||||||||||||||||||||||||||:::||||||||||||||||||||||||||||
Db        766 CGAGGAGAGTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAAC 825

Qy        277 TrpLysGlyGluAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArg 296
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        826 TGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGA 885

Qy        297 GlyIleCysAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyPro 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        886 GGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCT 945

Qy        317 GlyCysSerValProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsn 336
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        946 GGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAAC 1005

Qy        337 LeuLysLeuProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValVal 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1006 TTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTT 1065

Qy        357 GlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSer 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1066 GGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCT 1125

Qy        377 ArgGluTrpLeuProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSer 396
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1126 AGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCT 1185

Qy        397 LeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsn 416

Db       1186 TTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACTGGGAAT 1245

Qy        417 ValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrPro 436

Db       1246 GTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCT 1305

Qy        437 LysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsn 456

Db       1306 AAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAAT 1365

Qy        457 GlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsn 476

Db       1366 GGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAAT 1425

Qy        477 ValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeu 496

Db       1426 GTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTT 1485

Qy        497 ValGlnGlyGlyTyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrVal 516

Db       1486 GTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTT 1545

Qy        517 HisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArg 536

Db       1546 CATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGA 1605

Qy        537 TyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeu 556

Db       1606 TATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTG 1665

Qy        557 HisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAsp 576

Db       1666 CACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGAC 1725

Qy        577 ThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAla 596

Db       1726 ACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCC 1785

Qy        597 CysAspArgTrpSerValLeuProArgProAspLeuHisHisAspValAsnArgPheGly 616

Db       1786 TGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGC 1845

Qy        617 HisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeu 636

Db       1846 CATTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTC 1905

Qy        637 LeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAla 656

Db       1906 CTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCT 1965

Qy        657 CysLeuAlaAlaGlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIle 676

Db       1966 TGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATC 2025

Qy        677 SerTrpAlaLeuAlaThrAspGluGlnGluGluLysLeuLysSerGluCysPheSerLys 696

Db       2026 TCGTGGGCGCTGGCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAA 2085

Qy        697 ArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsn 716

Db       2086 AGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAAC 2145

Qy        717 ThrAsnAspCysHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGlu 736

Db       2146 ACCAATGACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAA 2205

Qy        737 GlyGlnIleSerIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCys 756

Db       2206 GGCCAGATCTCCATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGT 2265

Qy        757 AsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArg 776

Db       2266 AACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGG 2325

Qy        777 AsnGlnGluCysIleAlaLeuProGluAsnIleCys-GlyIleGlyTrp 792

Db       2326 AATCAGGAGTGCATTGCCCTGCCCGGTAGGCCTTGCAGGGTCATCTTGG 2374
```

RESULT 5
US-09-245-041-12
; Sequence 12, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 6370
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-245-041-12

Alignment Scores:

Pred. No.:	7.06e-306	Length:	6370
Score:	3491.50	Matches:	656
Percent Similarity:	95.51%	Conservative:	3
Best Local Similarity:	95.07%	Mismatches:	18
Query Match:	44.11%	Indels:	13
DB:	3	Gaps:	4

US-09-787-097-12 (1-1429) x US-09-245-041-12 (1-6370)

```
Qy        753 MetTyrTyrCysAsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGln 772

Db          1 ATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAG 60

Qy        773 TrpGluProArgAsnGlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGlyTrp 792

Db         61 TGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGG 120

Qy        793 HisLeuValGlyAsnSerCysLeuLysIleThrThrAlaLysGluAsnTyrAspAsnAla 812

Db        121 CATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCAAGGAGAATTATGACAATGCT 180

Qy        813 LysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSerLeuThrThrGlnLysLysVal 832

Db        181 AAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTA 240

Qy        833 GluPheValLeuLysGlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThr 852

Db        241 GAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACC 300

Qy        853 LeuThrProTrpValGly---LeuArgLysIleAsnValSerTyr-TrpCysTrp-GluA 871

Db        301 TTAACCCCATGGGTCGGGCCTTCGGGAAGGTYCAATGTGTCCTACTKGGTGCTGGGGAAG 360

Qy        871 spMetSerPro---PheThrAsnSer-LeuLeuGlnTrpMetProSerGluProSerAsp 889

Db        361 GATATGKTCCCCATTTTACAAATAGTTTTACTACAGTGGGATGSCCGTCTTGAGGCCCAG 420

Qy        890 Ala------GlyPheCysGly---IleLeuSerGluProSer--ThrArgGlyLeuLysA 906

Db        421 GCA------GGCTTCTGTGGC---ATCCTCTCTGAGCCCTCC--ACCAGAGGCCTCAAGA 469
```

```
Db        421 TGTTGCTTGGRATTCTGTGGGAATTTTATTCAGGAACCCAGTTACTTCGGGGACTGAAGG 480
Qy        906 laAlaThrCysIle-AsnProLeu-AsnGlySerValCysGluArgProAlaAsnHisSe 925
Db        481 CTGCAACCTGCATTCAACCCACTYMAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAG 540
Qy        925 rAlaLys-GlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThrSerG 945
Db        541 TGCTAAGGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCG 600
Qy        945 lySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrV 965
Db        601 GCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATG 660
Qy        965 alAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSerThrCysProProG 985
Db        661 TGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTG 720
Qy        985 luAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpC 1005
Db        721 AAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGT 780
Qy       1005 ysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyProV 1025
Db        781 GTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAG 840
Qy       1025 alLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerS 1045
Db        841 TGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCA 900
Qy       1045 erMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnC 1065
Db        901 GCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAAT 960
Qy       1065 ysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrT 1085
Db        961 GCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCA 1020
Qy       1085 hrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyGlyL 1105
Db       1021 CAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGA 1080
Qy       1105 ysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysC 1125
Db       1081 AATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGT 1140
Qy       1125 ysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGluAsnA 1145
Db       1141 GCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATC 1200
Qy       1145 rgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrGlnP 1165
Db       1201 GATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGT 1260
Qy       1165 heThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIleAsnPheValAlaT 1185
Db       1261 TCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTA 1320
Qy       1185 hrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnL 1205
Db       1321 CTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACC 1380
Qy       1205 euAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProV 1225
Db       1381 TCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTG 1440
Qy       1225 alValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPheAspP 1245
Db       1441 TTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATT 1500
Qy       1245 heArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpProIleL 1265
Db       1501 TTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCA 1560

Qy       1265 ysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePheValT 1285
Db       1561 AAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCTGGTACAGTTCTTCGTGA 1620
Qy       1285 hrPhePheSerCysPheLeuSerLeuLeuLeuValAlaAlaValValTrpLysIleLysG 1305
Db       1621 CTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGTGGTTTGGAAGATCAAAC 1680
Qy       1305 lnSerCysTrpAlaSerArgArgArgGluGlnLeuLeuArgGluMetGlnGlnMetAlaS 1325
Db       1681 AAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGAGATGCAACAGATGGCCA 1740
Qy       1325 erArgProPheAlaSerValAsnValAlaLeuGluThrAspGluGluProProAspLeuI 1345
Db       1741 GCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTA 1800
Qy       1345 leGlyGlySerIleLysThrValProLysProIleAlaLeuGluProCysPheGlyAsnL 1365
Db       1801 TTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGAGCCGTGTTTTGGCAACA 1860
Qy       1365 ysAlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyGlyIleProProP 1385
Db       1861 AAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCTGGGTGGCATCCCTCCTC 1920
Qy       1385 roGlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetProI 1405
Db       1921 CTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGA 1980
Qy       1405 leValTyrLysGluLysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnP 1425
Db       1981 TAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGC 2040
Qy       1425 roGlyThrCysIle 1429
Db       2041 CTGGGACCTGCATC 2054
```

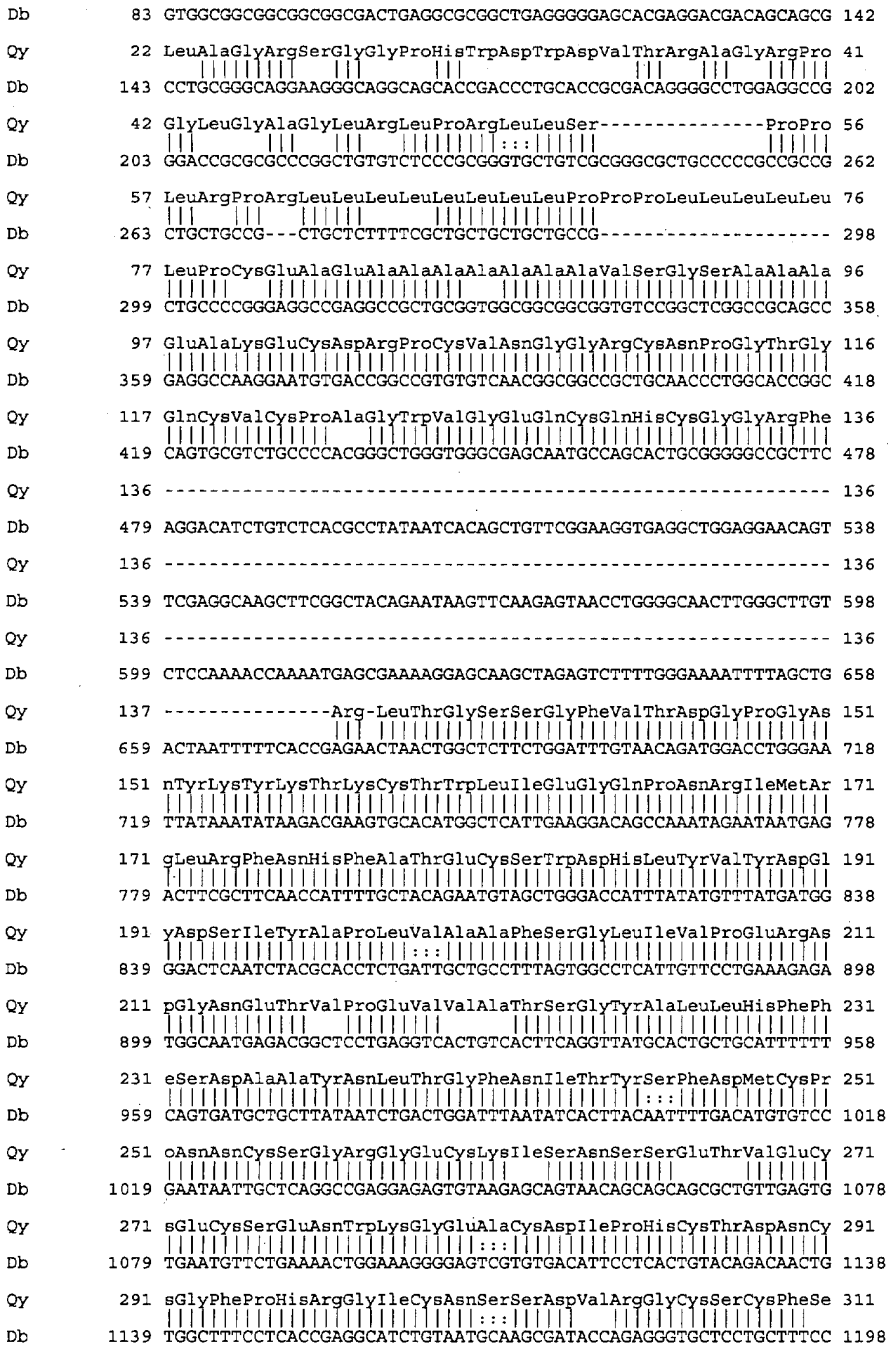
RESULT 6
US-09-245-041-8
; Sequence 8, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 2419
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-8

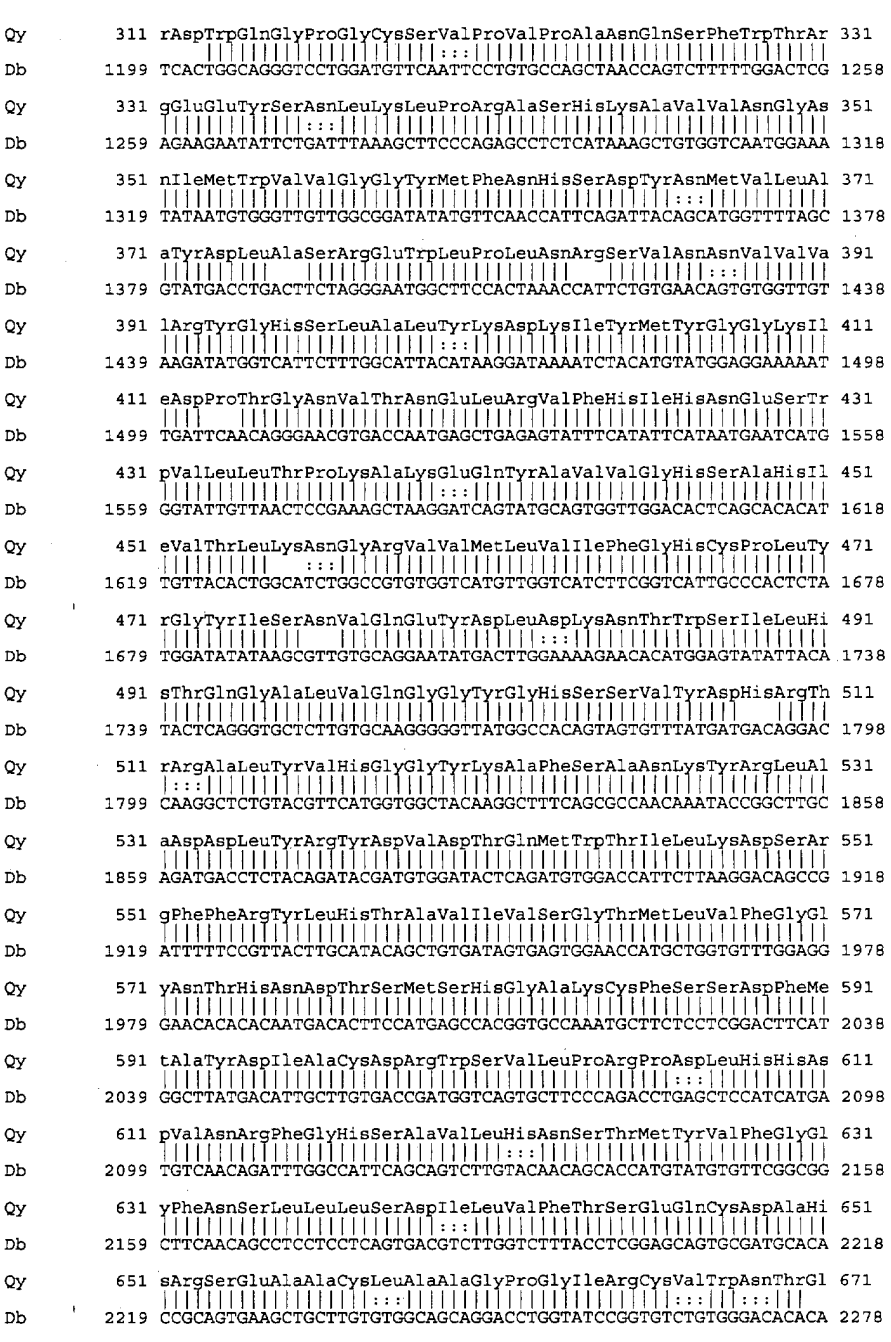
Alignment Scores:

Pred. No.:	1.71e-301	Length:	2419
Score:	3435.00	Matches:	642
Percent Similarity:	84.37%	Conservative:	22
Best Local Similarity:	81.58%	Mismatches:	44
Query Match:	43.40%	Indels:	79
DB:	3	Gaps:	4

US-09-787-097-12 (1-1429) x US-09-245-041-8 (1-2419)

```
Qy          2 ValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAlaAla 21
```





```
Qy        671 ySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnGluGluLysLeuLysSe 691
Db       2279 GTCGTCTCGATGTACCTCCTGGGAGTTGGCAACTGAAGAACAAGCAGAAAAGTTAAAATC 2338

Qy        691 rGluCysPheSerLysArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTy 711
Db       2339 AGAGTGTTTTTCTAAAAGAACCCTTGACCATGACAGATGTGACCAGCACACAGATTGTTA 2398

Qy        711 rSerCysThrAlaAsnThr 717
Db       2399 CAGCTGCACAGCCAATACC 2417

RESULT 7
US-09-245-041-10
; Sequence 10, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT: Nagle, D.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
;  TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
;  FILE REFERENCE: 7853-136
;  CURRENT APPLICATION NUMBER: US/09/245,041
;  CURRENT FILING DATE: 1999-02-05
;  EARLIER APPLICATION NUMBER: 60/093,630
;  EARLIER FILING DATE: 1998-07-21
;  EARLIER APPLICATION NUMBER: 60/104,978
;  EARLIER FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 131
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10
;   LENGTH: 1051
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-245-041-10

Alignment Scores:
Pred. No.:                1.52e-101      Length:          1051
Score:                    1226.50        Matches:         236
Percent Similarity:       84.45%         Conservative:    3
Best Local Similarity:    83.39%         Mismatches:      31
Query Match:              15.50%         Indels:          13
DB:                       3              Gaps:            3

US-09-787-097-12 (1-1429) x US-09-245-041-10 (1-1051)

Qy          2 ValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAlaAla 21
Db        217 GTGGCGGCGGCGGCGGCGACTGAGGCGCGGCTGAGGGGGAGCACGACGACGACAGCAGCG 276

Qy         22 LeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArgPro 41
Db        277 CCTGCGGGCAGGAAGGGCAGGCAGCACCGACCCTGCACCGCGACAGGGGCCTGGAGGCCG 336

Qy         42 GlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSer----------ProPro 56
Db        337 GGACCGCGCGCCCGGCTGTGTCTCCCGCGGGTGCTGTCGCGGGCGCTGCCCCCGCCGCCG 396

Qy         57 LeuArgProArgLeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeu 76
Db        397 CTGCTGCCG---CTGCTCTTTTCGCTGCTGCTGCTGCCG---------- 432

Qy         77 LeuProCysGluAlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAla 96
Db        433 CTGCCCCGGGAGGCCGAGGCCGCTGCGGTGGCGGCGGCGGTGTCCGGCTCGGCCGCAGCC 492

Qy         97 GluAlaLysGluCysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGly 116
Db        493 GAGGCCAAGGAATGTGACCGGCCGTGTGTCAACGGCGGCCGCTGCAACCCTGGCACCGGC 552

Qy        117 GlnCysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPhe 136
Db        553 CAGTGCGTCTGCCCCACGGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTC 612

Qy        137 ArgLeuThrGlySerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThr 156
Db        613 AGACTAACTGGCTCTTCTGGATTTGTAACAGATGGACCTGGGAATTATAAATATAAGACG 672

Qy        157 LysCysThrTrpLeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHis 176
Db        673 AAGTGCACATGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGCTTCAACCAT 732

Qy        177 PheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAla 196
Db        733 TTTGCTACAGAATGTAGCTGGGACCATTTATATGTTTATGATGGGGACTCAATCTACGCA 792

Qy        197 ProLeuValAlaAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrVal 216
Db        793 CCTCTGATTGCTGCCTTTAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCT 852

Qy        217 ProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyr 236
Db        853 CCTGAGGTCACTGTCACTTCAGGTTATGCACTGCTGCATTTTTTCAGTGATGCTGCTTAT 912

Qy        237 AsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGly 256
Db        913 AATCTGACTGGATTTAATATCACTTACAATTTTGACATGTGTCCGAATAATTGCTCAGGC 972

Qy        257 ArgGlyGluCysLysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsn 276
Db        973 CGAGGAGAGTGTAAGAGCAGTAACAGCAGCAGCGCTGTTGAGTGTGAATGTTCTGAAAAC 1032

Qy        277 TrpLysGly 279
Db       1033 TGGAAAGGG 1041

RESULT 8
US-09-245-041-5
; Sequence 5, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT: Nagle, D.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
;  TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
;  FILE REFERENCE: 7853-136
;  CURRENT APPLICATION NUMBER: US/09/245,041
;  CURRENT FILING DATE: 1999-02-05
;  EARLIER APPLICATION NUMBER: 60/093,630
;  EARLIER FILING DATE: 1998-07-21
;  EARLIER APPLICATION NUMBER: 60/104,978
;  EARLIER FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 131
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5
;   LENGTH: 90050
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-245-041-5

Alignment Scores:
Pred. No.:                2.73e-56       Length:          90050
Score:                    766.00         Matches:         205
Percent Similarity:       21.41%         Conservative:    2
Best Local Similarity:    21.20%         Mismatches:      6
Query Match:              9.68%          Indels:          755
DB:                       3              Gaps:            2

US-09-787-097-12 (1-1429) x US-09-245-041-5 (1-90050)

Qy        849 SerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868
Db      28996 TCCAAGCTCACTCTGACTCCATGGGTTGGTCTTCGGAAGATCAATGTGTCTTACTGGTGC 29055

Qy        869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888
```

```
Db      29056 TGGGAGGATATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGT 29115
Qy        889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      29116 GATGCTGGCTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGATTAAAGGCTGCAACC 29175
Qy        909 CysIleAsnProLeuAsnGlySerValCysGluArgProAla---------------- 922
              |||||||||||||||||||||||||||||||||||||||
Db      29176 TGCATCAACCCTCTCAATGGCAGCGTCTGTGAAAGGCCTGG-TAAGGACATGGGTGCATA 29234
Qy        922 ---------------------------------------------------------- 922
Db      29235 TAGTGCTCCAGGAGGAGCCAAGACAGCAAAGGAGGCACAGCTGAATGAGCGCTGAGGTGA 29294
Qy        922 ---------------------------------------------------------- 922
Db      29295 TGAAGTACTTATGGCAGCAGGGAGAGGAGCACCAATTTAGGCATATGTATTTCAAACAGA 29354
Qy        922 ---------------------------------------------------------- 922
Db      29355 ACCCGATTCCAGATAGTCTTTCTTGGCCTCTGACTGCTTTAAGCCATACTGAAAACCAAA 29414
Qy        922 ---------------------------------------------------------- 922
Db      29415 AATAAAATTGCTGAAAGAACCCAGTTTATATTGAGCTGCACTGTTTCGTTGGTCTCAAAG 29474
Qy        922 ---------------------------------------------------------- 922
Db      29475 TGTTGAGAATTGTTCTAGAAGATTATTTCCTTGGTGTTGGCAGAGAAGTGCTATGGAGGA 29534
Qy        922 ---------------------------------------------------------- 922
Db      29535 AACAACAACCTGAAACCAAAGAAACATTTAGAAAAGCAGCAAGTCAGGACACTATTCAGA 29594
Qy        922 ---------------------------------------------------------- 922
Db      29595 CACTGCTGGGGTGGGGGGAGAGGGGCATGGCCAAAGAAGCCGACAGAGCCAACACCAGGC 29654
Qy        922 ---------------------------------------------------------- 922
Db      29655 TGTGGCAATGTCCTGCGCTGAGGTTAAGGTTAGACTCCATGAGGCCAGGCCCAGAACAGC 29714
Qy        922 ---------------------------------------------------------- 922
Db      29715 CATACACAAATGAGGACTCCAAAACAAGAGGTGCAAGTGTAGTGGAGACTCCATCCCTGC 29774
Qy        922 ---------------------------------------------------------- 922
Db      29775 AGGTCCTGTTTCAGGAAATGATTGTACTTTGCCTGAGTAATACAGCCTAGGAGCTACTTT 29834
Qy        922 ---------------------------------------------------------- 922
Db      29835 CTGATAGGGTTTTTTAAATACTTACAAAGAATTATTTATCTTTAATCATGTGGTTTTGTA 29894
Qy        922 ---------------------------------------------------------- 922
Db      29895 TGTGTGTGCTTGCACATGCAGTGCTTGTGAGAGAGAGTATGTGTGAGAGCATGCATGTAT 29954
Qy        922 ---------------------------------------------------------- 922
Db      29955 GAGAGTGTGAGAATATATGTGAGAGAGTGTGAGTGCATGTGTGCGTGTGTGCATCTGTGT 30014
Qy        922 ---------------------------------------------------------- 922
Db      30015 GTACAGGTGTGTGTACATGCATGTGTGTATAAGAGTATGTGAGAGTGTGGGTGTGTGTGT 30074
Qy        922 ---------------------------------------------------------- 922
Db      30075 GAGAGTATGTGAGAATATATGTATGAGTGTGTGTGAGTATGAGTGTATGTGCGTGCCTGC 30134
Qy        922 ---------------------------------------------------------- 922
Db      30135 ATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTAGAAGTGGCCTTGG 30194

Qy        922 ---------------------------------------------------------- 922
Db      30195 AAAACAGAGTTGTCAGATCTCTTAGAGATATAGTTGCAGTTGGTTGTGAGCCATCTCATA 30254
Qy        922 ---------------------------------------------------------- 922
Db      30255 TGAGCGCTGGAAGTTGAAATTGGGTTCTCTGGAATCCTCTGGGTTCCTTGTTGAAGCCTG 30314
Qy        922 ---------------------------------------------------------- 922
Db      30315 AATATTTTGATAAATATTTATGTCATTATCCCTCAAAATTGTAAATGTAGAATTTAACAA 30374
Qy        922 ---------------------------------------------------------- 922
Db      30375 ACTCAGGTCTTGAGTCATCTTTGTCCCAAGGTTTGTTTGTTTGGTTTTTTGTTCCCCCAC 30434
Qy        922 ---------------------------------------------------------- 922
Db      30435 CTTTTCTTCAGTGCTTTTAAAAAAGAGAGTCCATTTTTTCCTAAATGTTTAAATACAGTT 30494
Qy        922 ---------------------------------------------------------- 922
Db      30495 GAGGAATAGAACATCTGACTCCAATTTCCTGGGTTTCCCTCCATGTAGTGTAGTGCTGAC 30554
Qy        922 ---------------------------------------------------------- 922
Db      30555 CTGATTTCAGTGTGCATTGAAAACTTTGATCACTTGGAAGGCAGCTATGCTCACCACTAT 30614
Qy        922 ---------------------------------------------------------- 922
Db      30615 ACTACCAATGTCTGCAATCCTATAGGAGAAACAACAATATGAACTAACTAGTACCCCCCA 30674
Qy        922 ---------------------------------------------------------- 922
Db      30675 GAGCTGTGTCTCTAGTTGCATATGTAGCAGAGGATGGCCTAGTCAGCCATCATTGGGAGG 30734
Qy        922 ---------------------------------------------------------- 922
Db      30735 AGAGGCCCTTGGTATTGCGAAGATCATATGCCCCAGTACAGGGGAATGCCAGGACCAGGA 30794
Qy        922 ---------------------------------------------------------- 922
Db      30795 AGCAAGAGTGGGTGGGTTGGGGAGCAGTGCGGGGGGGGGGGGGTATAGGGGGTTTTGGGGA 30854
Qy        922 ---------------------------------------------------------- 922
Db      30855 TAGCATTTGAAATGTAAATGAAGAAAATAACTAATAAAAATTGCCTTAAAAAAAAACAAA 30914
Qy        923 ----------------------------------------------------------Asn 923
                                                                         |||
Db      30915 AAAGAAAAGTTTTTGATCTTAGCTGACCAGTGTCTCTTTGGGTCTTAATTTCCAGCAAAC 30974
Qy        924 HisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThr 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||
Db      30975 CACAGTGCCAAGCAGTGCCGGACACCATGTGCCCTGCGGACAGCGTGTGGCGAGTGCACT 31034
Qy        944 SerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAla 963
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      31035 AGCAGCAGCTCGGAGTGCATGTGGTGCAGTAACATGAAGCAGTGTGTGGACTCCAATGCC 31094
Qy        964 TyrValAlaSerProProPheGlyGlnCysMetGluTrpTyrThrMetSerThrCysPro 983
              ||||||||||||||||||||||||||||||||||||||||||||||||||:::||||||
Db      31095 TACGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCAGCTGCCCA 31154
Qy        983 ---------------------------------------------------------- 983
Db      31155 CGTAAGTGGAAGGAGCTTTTGAACATTTGCAGGCAAGTTGGGCTTGACTTTCTGCTCAAG 31214
Qy        983 ---------------------------------------------------------- 983
Db      31215 TCCATGCAGAAGCTGGTCGGGCCGGCCCTTCCAGATTAACATGTATGTATAGAATGCAGC 31274
```

```
Qy        983 ------------------------------------------------------------ 983

Db      31275 ACAGTGTTCCATGCAGTAAATCAGTTACATCAAGGAGAAGGCACAGGGTACAGAAATACC 31334

Qy        983 ------------------------------------------------------------ 983

Db      31335 TTTTCTTCTTCAGGGTAATATTATAATTCAATCTGTATAATGTTTCTACATCTTAATCTA 31394

Qy        983 ------------------------------------------------------------ 983

Db      31395 CCAGTATGTAAAGTGCTTTCTAGTAGAGGCCTCCCCAGCTCCCTTTTTCATCCAACATCC 31454

Qy        983 ------------------------------------------------------------ 983

Db      31455 TGATATTAAAAGGTTGGAAAAGTCCCTGTTATATATTATGTAAAATGTGGGGCCCTTTAA 31514

Qy        983 ------------------------------------------------------------ 983

Db      31515 ATTATTTCAGTTCAATAATCACTATAGGGTACTATTTTTAATTCATGGAAGTTAAATCAT 31574

Qy        983 ------------------------------------------------------------ 983

Db      31575 CTGTTAAAAGAAAAGGTAATAACAGTAAATTCAAATCTTGTGATAGTGAATTACAAGTTG 31634

Qy        984 -------------------------ProGluAsnCysSerGlyTyrCysThrCysSerHi 995
                                        ||||||||||||||||||||||||||||||||
Db      31635 GATTGTTTTGCCTTGTTTTTTAATAGCTGAAAATTGCTCTGGCTACTGTACCTGCAGCCA 31694

Qy        995 sCysLeuGluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLy 1015
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      31695 TTGCTTGGAGCAGCCAGGCTGTGGTTGGTGTACTGATCCTAGCAATACTGGGAAAGGAAA 31754

Qy       1015 sCysIleGluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAs 1035
              ||||||||||||||||||||||||||||||||||||||||||||||||||||      |||||
Db      31755 ATGTATTGAGGGCAGCTATAAAGGACCTGTGAAGATGCCGTCACAGGCCTCTGCAGGAAA 31814

Qy       1035 nPheTyrProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTr 1055
              |   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      31815 TGTGTATCCACAGCCCCTTCTGAACTCCAGCATGTGTCTAGAGGACAGCAGATACAACTG 31874

Qy       1055 pSerPheIleHisCysPro 1061
              |||||||||||||||||
Db      31875 GTCTTTCATTCACTGTCCA 31893

RESULT 9
US-09-833-381-1916
; Sequence 1916, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
;  APPLICANT: Robison, Keith E.
;  TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
;  FILE REFERENCE: 5800-119
;  CURRENT APPLICATION NUMBER: US/09/833,381
;  CURRENT FILING DATE:  2001-04-11
;  PRIOR APPLICATION NUMBER: 09/516,448
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 2050
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1916
;   LENGTH: 415
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(415)
;   OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1916

Alignment Scores:
Pred. No.:              1.11e-43        Length:         415
Score:                  581.50          Matches:        105
Percent Similarity:     84.56%          Conservative:   10
Best Local Similarity:  77.21%          Mismatches:     20
Query Match:            7.35%           Indels:         1
DB:                     4               Gaps:           1

US-09-787-097-12 (1-1429) x US-09-833-381-1916 (1-415)

Qy        141 SerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrp 160
              :::||||||::::::||||||||||||   ||||||||||||||||||||||||||||
Db          4 GCGTCCGGATATTTAACAGATGGCCCAATTAACTATAAATATAAAACTAAATGTACTTGG 63

Qy        161 LeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGlu 180
              ||||||||||||   ||||||   ::::::||||||||||||||||||||||||||||
Db         64 CTCATTGAAGGCTATCCAAATGCAGTGTTAAGATTAAGATTCAATCATTTTGCTACAGAA 123

Qy        181 CysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAla 200
              ||||||||||||||:::|||||||||||||||||||||||||||||||||||:::|||
Db        124 TGTAGCTGGGATCATATGTATGTTTATGATGGAGATTCAATATATGCACCTTTAATAGCT 183

Qy        201 AlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGluValVal 220
              ||||||||||||||||||||||      ||||||||||||||||||||||||||||
Db        184 GTACTTAGTGGTTTGATAGTCCCTGAAATAAGGGGCAATGAAACTGTGCCTGAAGTTGTT 243

Qy        221 AlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGly 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 ACTACATCTGGCTATGCACTGTTACATTTTTTTAGTGATGCTGCGTATAATCTAACTGGT 303

Qy        241 PheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCys 260
              |||||||||  ||||||  :::   ||||||||||||||||||||  |||:::|||
Db        304 TTCAACATTTTCTATTCAATCAATTCTTGTCCTAACAATTGCTCTGGTCATGGGAAGTGT 363

Qy        261 LysIleSerAsnSer---SerGluThrValGluCysGluCysSerGlu 275
                  |||   |||           |||   |||||||||   :::
Db        364 ACAACTAGTGTCTCTGTTCCAAGTCAAGTATATTGTGAATGTGATAAA 411

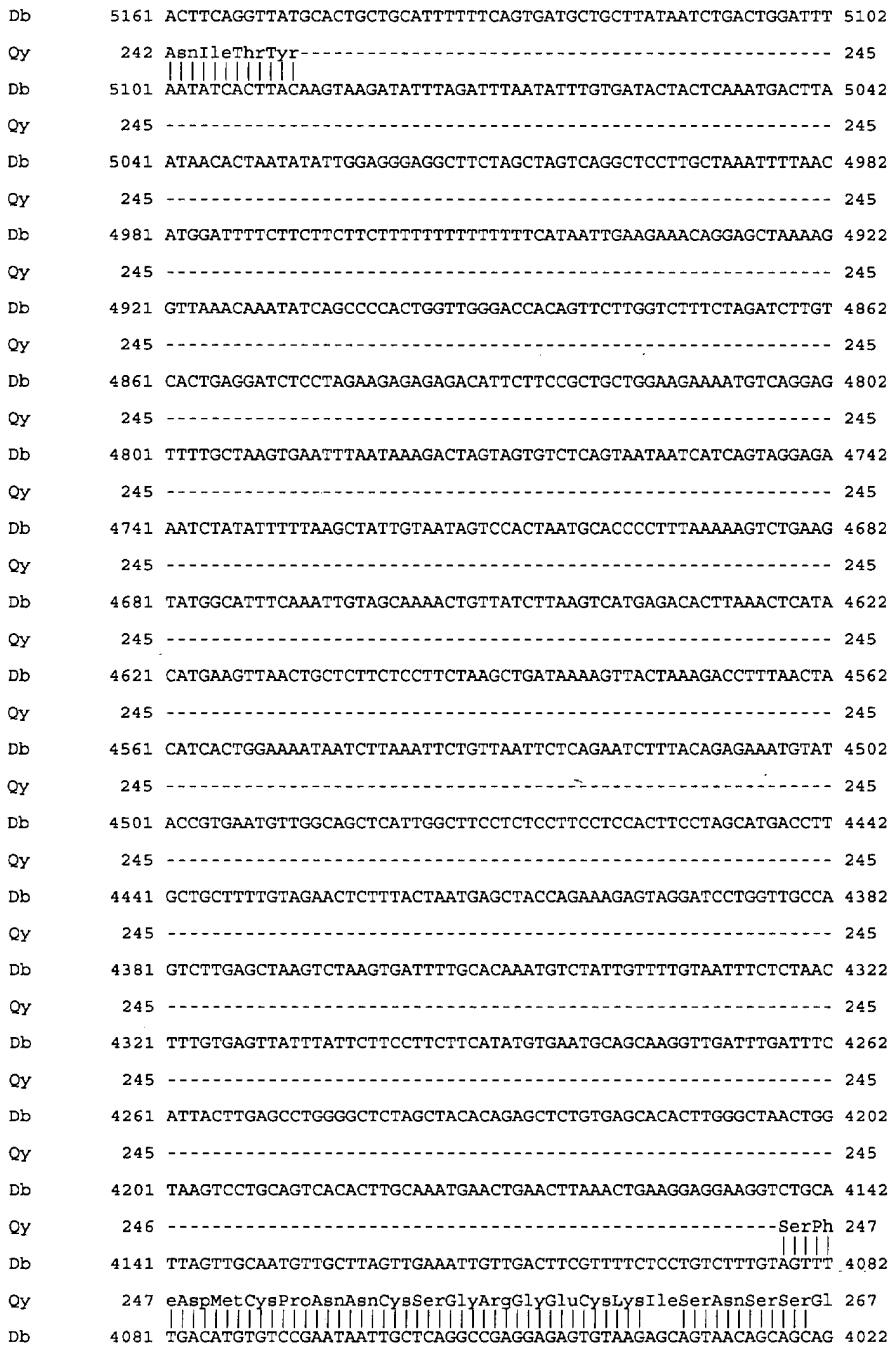
RESULT 10
US-09-245-041-3/c
; Sequence 3, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT: Nagle, D.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
;  TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
;  FILE REFERENCE: 7853-136
;  CURRENT APPLICATION NUMBER: US/09/245,041
;  CURRENT FILING DATE: 1999-02-05
;  EARLIER APPLICATION NUMBER: 60/093,630
;  EARLIER FILING DATE: 1998-07-21
;  EARLIER APPLICATION NUMBER: 60/104,978
;  EARLIER FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 131
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 17056
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-245-041-3

Alignment Scores:
Pred. No.:              3.1e-29         Length:         17056
Score:                  454.00          Matches:        155
Percent Similarity:     16.81%          Conservative:   5
Best Local Similarity:  16.28%          Mismatches:     10
Query Match:            5.74%           Indels:         783
DB:                     3               Gaps:           2

US-09-787-097-12 (1-1429) x US-09-245-041-3 (1-17056)

Qy        202 PheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGluValValAla 221
              ||||||||||||||||||||||||||||||||||||||||||||    |||||||||
Db       5221 TTCAGTGGCCTCATTGTTCCTGAAAGAGATGGCAATGAGACGGCTCCTGAGGTCACTGTC 5162

Qy        222 ThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPhe 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```


Db 5161 ACTTCAGGTTATGCACTGCTGCATTTTTTCAGTGATGCTGCTTATAATCTGACTGGATTT 5102
Qy 242 AsnIleThrTyr 245
Db 5101 AATATCACTTACAAGTAAGATATTTAGATTTAATATTTGTGATACTACTCAAATGACTTA 5042
Qy 245 245
Db 5041 ATAACACTAATATATTGGAGGGAGGCTTCTAGCTAGTCAGGCTCCTTGCTAAATTTTAAC 4982
Qy 245 245
Db 4981 ATGGATTTTCTTCTTCTTCTTTTTTTTTTTTTTCATAATTGAAGAAACAGGAGCTAAAAG 4922
Qy 245 245
Db 4921 GTTAAACAAATATCAGCCCCACTGGTTGGGACCACAGTTCTTGGTCTTTCTAGATCTTGT 4862
Qy 245 245
Db 4861 CACTGAGGATCTCCTAGAAGAGAGAGACATTCTTCCGCTGCTGGAAGAAAATGTCAGGAG 4802
Qy 245 245
Db 4801 TTTTGCTAAGTGAATTTAATAAAGACTAGTAGTGTCTCAGTAATAATCATCAGTAGGAGA 4742
Qy 245 245
Db 4741 AATCTATATTTTTAAGCTATTGTAATAGTCCACTAATGCACCCCTTTAAAAAGTCTGAAG 4682
Qy 245 245
Db 4681 TATGGCATTTCAAATTGTAGCAAAACTGTTATCTTAAGTCATGAGACACTTAAACTCATA 4622
Qy 245 245
Db 4621 CATGAAGTTAACTGCTCTTCTCCTTCTAAGCTGATAAAAGTTACTAAAGACCTTTAACTA 4562
Qy 245 245
Db 4561 CATCACTGGAAAATAATCTTAAATTCTGTTAATTCTCAGAATCTTTACAGAGAAATGTAT 4502
Qy 245 245
Db 4501 ACCGTGAATGTTGGCAGCTCATTGGCTTCCTCTCCTTCCTCCACTTCCTAGCATGACCTT 4442
Qy 245 245
Db 4441 GCTGCTTTTGTAGAACTCTTTACTAATGAGCTACCAGAAAGAGTAGGATCCTGGTTGCCA 4382
Qy 245 245
Db 4381 GTCTTGAGCTAAGTCTAAGTGATTTTGCACAAATGTCTATTGTTTTGTAATTTCTCTAAC 4322
Qy 245 245
Db 4321 TTTGTGAGTTATTTATTCTTCCTTCTTCATATGTGAATGCAGCAAGGTTGATTTGATTTC 4262
Qy 245 245
Db 4261 ATTACTTGAGCCTGGGGCTCTAGCTACACAGAGCTCTGTGAGCACACTTGGGCTAACTGG 4202
Qy 245 245
Db 4201 TAAGTCCTGCAGTCACACTTGCAAATGAACTGAACTTAAACTGAAGGAGGAAGGTCTGCA 4142
Qy 246 SerPh 247
Db 4141 TTAGTTGCAATGTTGCTTAGTTGAAATTGTTGACTTCGTTTTCTCCTGTCTTTGTAGTTT 4082
Qy 247 eAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSerSerGl 267
Db 4081 TGACATGTGTCCGAATAATTGCTCAGGCCGAGGAGAGTGTAAGAGCAGTAACAGCAGCAG 4022
Qy 267 uThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIleProHisCy 287
Db 4021 CGCTGTTGAGTGTGAATGTTCTGAAAACTGGAAAGGGGAGTCGTGTGACATTCCTCACTG 3962
Qy 287 sThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspValArgGlyCy 307
Db 3961 TACAGACAACTGTGGCTTTCCTCACCGAGGCATCTGTAATGCAAGCGATACCAGAGGGTG 3902
Qy 307 sSerCysPheSerAspTrpGln 314
Db 3901 CTCCTGCTTTCCTCACTGGCA-GGGTAGGAGCTGCCCTTCAAAGTCCTTTCCCTCACAGA 3843
Qy 314 314
Db 3842 AAGTTCTGAGTTAGTGAACTGTAGGGAGTTTATTTTAGCAGCCATACGCCTTATGTTAAT 3783
Qy 314 314
Db 3782 CATGTATCAGGTGGCATTTAAGCATTCATCTCCATTTATTCAGATTCTTTGAATTTACAT 3723
Qy 314 314
Db 3722 ATATGTGGTTGTTATCTATGTGCTTAGAATTGTGGAAAGGTTTGACAAATCCTTTGCAAC 3663
Qy 314 314
Db 3662 TCATTCGTGGGAATTAGGGAACATGGTCTGGTGATAGAGAAGCAGAGACTCAAGCTGAGA 3603
Qy 314 314
Db 3602 GGAGGACAAGCGTTCCTTAGAGTCTCAGTGCTGTCATCTAGTGCCATTACCCAACACAAC 3543
Qy 314 314
Db 3542 CGGGTACTAAGGTACTGTAGGGAATTCCCTGGGGTATTGAGGAAGCAGTCATCGTTGGTT 3483
Qy 314 314
Db 3482 TAGTGATTTTCCTTCTATGCTTAGGTCCTCAGAAAGTAAGTGTTCATGTAAATGTTTTTA 3423
Qy 314 314
Db 3422 GTTTGTCCAATTCTCCCTTTGTCTCTTTCATCTATTCAGCGCCTGGACTTGCCAGAAGAT 3363
Qy 314 314
Db 3362 CGATTTTGATCTGTCATTATATTTCAACCCTAATGCCTTTTCCACATGAGTTAATATAAT 3303
Qy 314 314
Db 3302 AACCCCTTGGTCTACCCTTTCTACTAAATGTTGCTTCAGCAAGATTTTTACCTCTAAAAA 3243
Qy 314 314
Db 3242 AGTTTAACCTGATTATCAGAAATTGCCAAATTATCAGAAAGTGTGTAGTATTAAATTTTA 3183
Qy 314 314
Db 3182 TAAGGGAGTGATTGGTTTTGGTCTGTATCTTGAGTAAGTTAAAAACTCATGTTAGGTGCT 3123
Qy 314 314
Db 3122 AAGCAGATGGCTTCTTGGTTAAGAGTACTGGTTGCTCTTCCAGAGGACTCAGGTTTGATC 3063
Qy 314 314
Db 3062 CTAGCATCCACAAAGTAGCTCATAAGATCTGTAACTCTAGTTCCAGGGGATTCAATGCCT 3003
Qy 314 314
Db 3002 TTTCCTGACATCATTGGGTACCAAGCAAGCAAGTGGTATATAGGCATACATGTAAACAAA 2943

```
Qy        314 ------------------------------------------------------------ 314

Db       2942 ATACCTGTACATGTAAAAACCTTTAAAACCCCACAAGTTAAAGATGTTTGCGCTTAAGTG 2883

Qy        314 ------------------------------------------------------------ 314

Db       2882 CTAAAGGAAACACAAGTCTTTTTTTTTTTTTTTTCTGGTCTGTTTTTTAAAAAATTAA 2823

Qy        314 ------------------------------------------------------------ 314

Db       2822 TTTATTTATTTCATGTGAGTACATTGTTGCTGTCTTCAGACACACCAGAAGAGGGCATCA 2763

Qy        314 ------------------------------------------------------------ 314

Db       2762 GATCCCCATTACATATGGTTGTGAGCCACCATGTAGTTGCTGGGATTTGAACTCAGGACC 2703

Qy        314 ------------------------------------------------------------ 314

Db       2702 TCCAAGAAGAGCAGTCAGTTTTCTTAACCACTGAGCCATCTCTCTAGCCCCACAAGTCTT 2643

Qy        314 ------------------------------------------------------------ 314

Db       2642 TAAATATTCTTTGCACTTAGATTGGAACTTGGAATACTATAAACTTTCAGGCTGCATTTT 2583

Qy        315 ----------------------------------------GlyProGlyCysSerVa 320
                                                      ||||||||||||||::
Db       2582 GTGAAGCTATTTGTTGTTGTTGTTTTTTTTTTTTTTATTAGGTCCTGGATGCTCAAT 2523

Qy        320 lProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuPr 340
               :|||||||||||||||||||||||||||||||||||||||||||:::|||||||||||
Db       2522 TCCTGTGCCAGCTAACCAGTCTTTTGGACTCGAGAAGAATATTCTGATTTAAAGCTTCC 2463

Qy        340 oArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMe 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2462 CAGAGCCTCTCATAAAGCTGTGGTCAATGGAAATATAATGTGGGTTGTTGGCGGATATAT 2403

Qy        360 tPheAsnHisSerAspTyrAsnMetValLeuAla 371
               ||||||||||||||||||:::||||||||||||
Db       2402 GTTCAACCATTCAGATTACAGCATGGTTTTAGCG 2369
```

RESULT 11
US-09-702-705-966/c
; Sequence 966, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-966

Alignment Scores:

Pred. No.:	8.93e-32	Length:	246
Score:	446.00	Matches:	82
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.63%	Indels:	0
DB:	4	Gaps:	0

US-09-787-097-12 (1-1429) x US-09-702-705-966 (1-246)

```
Qy        839 LeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGly 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy        859 LeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSer 878
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy        879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGlu 898
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67

Qy        899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy        919 GluArg 920
              ||||||
Db          6 GAAAGG 1
```

RESULT 12
US-09-736-457-966/c
; Sequence 966, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-966

Alignment Scores:

Pred. No.:	8.93e-32	Length:	246
Score:	446.00	Matches:	82
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.63%	Indels:	0
DB:	4	Gaps:	0

US-09-787-097-12 (1-1429) x US-09-736-457-966 (1-246)

```
Qy        839 LeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGly 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy        859 LeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSer 878
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy        879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGlu 898
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67
```

```
Qy        899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy        919 GluArg 920
              ||||||
Db          6 GAAAGG 1

RESULT 13
US-09-614-124B-966/c
; Sequence 966, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
;  APPLICANT: Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.478C9
;  CURRENT APPLICATION NUMBER: US/09/614,124B
;  CURRENT FILING DATE:  2001-07-11
;  NUMBER OF SEQ ID NOS: 1668
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 966
;   LENGTH: 246
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-614-124B-966

Alignment Scores:
Pred. No.:                8.93e-32        Length:         246
Score:                    446.00          Matches:        82
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              5.63%           Indels:         0
DB:                       4               Gaps:           0

US-09-787-097-12 (1-1429) x US-09-614-124B-966 (1-246)

Qy        839 LeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGly 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy        859 LeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSer 878
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy        879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGlu 898
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67

Qy        899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy        919 GluArg 920
              ||||||
Db          6 GAAAGG 1

RESULT 14
US-09-671-325-966/c
; Sequence 966, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Fan, Liqun
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.478C12
;  CURRENT APPLICATION NUMBER: US/09/671,325
;  CURRENT FILING DATE:  2000-09-26
;  NUMBER OF SEQ ID NOS:  1825
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 966
;   LENGTH: 246
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-671-325-966

Alignment Scores:
Pred. No.:                8.93e-32        Length:         246
Score:                    446.00          Matches:        82
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              5.63%           Indels:         0
DB:                       4               Gaps:           0

US-09-787-097-12 (1-1429) x US-09-671-325-966 (1-246)

Qy        839 LeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGly 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy        859 LeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSer 878
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy        879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGlu 898
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67

Qy        899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy        919 GluArg 920
              ||||||
Db          6 GAAAGG 1

RESULT 15
US-08-083-590A-21
; Sequence 21, Application US/08083590A
; Patent No. 5786158
;  GENERAL INFORMATION:
;    APPLICANT:  Artavanis-Tsakonas, S. et al.
;    TITLE OF INVENTION:  Therapeutic And Diagnostic Methods
;    TITLE OF INVENTION:  And Compositions Based On No. 5786158ch Proteins And
;    TITLE OF INVENTION:  Nucleic Acids
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/083,590A
```

```
;      FILING DATE:  25-JUN-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Misrock, S. Leslie
;      REGISTRATION NUMBER:  18,872
;      REFERENCE/DOCKET NUMBER:  7326-015
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212 790-9090
;      TELEFAX:  212 8698864/9741
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9723 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  10..7419
US-08-083-590A-21

Alignment Scores:
Pred. No.:              8.2e-21       Length:          9723
Score:                  356.50        Matches:         275
Percent Similarity:     30.85%        Conservative:    139
Best Local Similarity:  20.49%        Mismatches:      478
Query Match:            4.50%         Indels:          454
DB:                     1             Gaps:            67

US-09-787-097-12 (1-1429) x US-08-083-590A-21 (1-9723)

Qy         53 LeuSerProProLeuArgProArgLeuLeuLeuLeuLeuLeuLeuProProProLeu 72
Db          4 ATTCCGCCCGCCCTGCGCCCCGCTCTGCTGTGG--------------------GCGCTG 42

Qy         73 LeuLeuLeuLeuLeuProCysGluAlaGluAlaAlaAlaAlaAlaAlaAlaValSerGly 92
Db         43 CTGGCGCTCTGGCTGTGCTGCGCGGCCCCCGCG-------------------------- 75

Qy         93 SerAlaAlaAlaGluAlaLysGluCysAspArgProCysValAsnGlyGlyArgCys--- 111
Db         76 ---CATGCATTGCAGTGTCGAGATGGCTATGAACCCTGTGTAAATGAAGGAATGTGTGTT 132

Qy        112 ------AsnProGlyThrGlyGlnCysValCysProAlaGlyTrpValGlyGluGlnCys 129
Db        133 ACCTACCACAATGGCACAGGATACTGCAAATGTCCAGAAGGCTTCTTGGGGGAATATTGT 192

Qy        130 GlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPheValThrAspGlyPro 149
Db        193 CAACATCGAGACCCCTGTGAGAAGAACCGCTGCCAGAATGGTGGGACTTGTGT-GGCCCA 251

Qy        150 GlyAsnTyrLysTyrLysThrLysCysThrTrpLeuIleGluGlyGlnProAsnArgIle 169
Db        252 GGCCAT---------------------------------GCTGGGGAAAGCCACGTGCCG 278

Qy        170 MetArgLeuArg----------------------PheAsnHisPheAlaThrGluCys 181
Db        279 ATGTGCCTCAGGGTTTACAGGAGAGGACTGCCAGTACTCAACATCTCATCCATGCTTTGT 338

Qy        182 --------------SerTrpAspHisLeu-TyrValTyrAspGlyAspSerIleTyrAl 196
Db        339 GTCTCGACCCTGCCTGAATGGCGGCACATGCCATATGCTCAGCCGGGATACCTATGAGTG 398

Qy        196 aProLeuValAlaAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrVa 216
Db        399 CACCTGTCAAGTCGGGTTTACAGGTAAGGAGTGCCAATGGACGGATGCCTGCCTGTCTCA 458

Qy        216 lProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTy 236
Db        459 TCCCTGTGCAAATGGAAGTACCTGTACCACTGTGGCCAACCAGTTCTCCTGCAAATGC-- 516

Qy        236 rAsnLeuThrGlyPhe---------AsnIleThrTyrSerPheAspMetCys------Pr 251

Db        517 ----CTCACAGGCTTCACAGGGCAGAAATGTGAGACTGATGTCAATGAGTGTGACATTCC 572

Qy        251 oAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSerSerGluThrValGluCy 271
Db        573 AGGACACTGCCAGCATGGTGGCACCTGC------CTCAACCTGCCTGGTTCCTACCAGTG 626

Qy        271 sGluCysSerGluAsnTrpLysGlyGluAlaCysAsp---------IleProHisCysTh 288
Db        627 CCAGTGCCCTCAGGGCTTCACAGGCCAGTACTGTGACAGCCTGTATGTGCCCTGTGCACC 686

Qy        288 rAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAsp------ValArgGl 306
Db        687 CTCACCTTGT------GTCAATGGAGGCACCTGTCGGCAGACTGGTGACTTCACTTTTGA 740

Qy        306 yCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAlaAsnGl 326
Db        741 GTGCAACTGCCTTCCAGGTTTTGAAGGGAGCACCTGTGAGAGGAATATT----------- 789

Qy        326 nSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHisLysAl 346
Db        790 ---------------GATGACTGCCCTAAC--------------------CACAGGTG 812

Qy        346 aValValAsnGlyAsnIle-MetTrpVal--------------ValGlyGlyTyrMetP 361
Db        813 T---CAGAATGGAGGGGTTTGTGTGGATGGGGTCAACACTTACAACTGCCGCTGTCCCCC 869

Qy        361 heAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuP 381
Db        870 ACAATGGACAGGACAGTTCT--------------------GCACAGAGGATGTGG---- 904

Qy        381 roLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrL 401
Db        905 --ATGAATGCCTGCTGCAGCCCAATGCCTGTCAAAATGGGGGCACCTGTGCCAACCGCAA 962

Qy        401 ysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsnValThrAsnGluL 421
Db        963 TGGAGGCTATGGCTGTGTAT---------GTGTCAACGGCTGGAGTGGAGATGACTGCAG 1013

Qy        421 euArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluG 441
Db       1014 TGAGAACATTGATGATTGTGCCTTCGCCTCCTGTACTCCAGGCTCCACCTGCATCGACCG 1073

Qy        441 lnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValM 461
Db       1074 TGTGG--------------CCTCCTTCTCTTGCATGTGCCCAGAGGGGAAGGCAGGTCT 1118

Qy        461 etLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrA 481
Db       1119 CCTGTGTCATC---------------------------------------------------- 1129

Qy        481 spLeuAspLysAsnThrTrpSerIleLeuHisThr-GlnGlyAlaLeuValGlnGlyGly 500
Db       1130 ----------------TGG---ATGATGCATGCATCAGCAATCCTTGCCACAAGGGGGC 1169

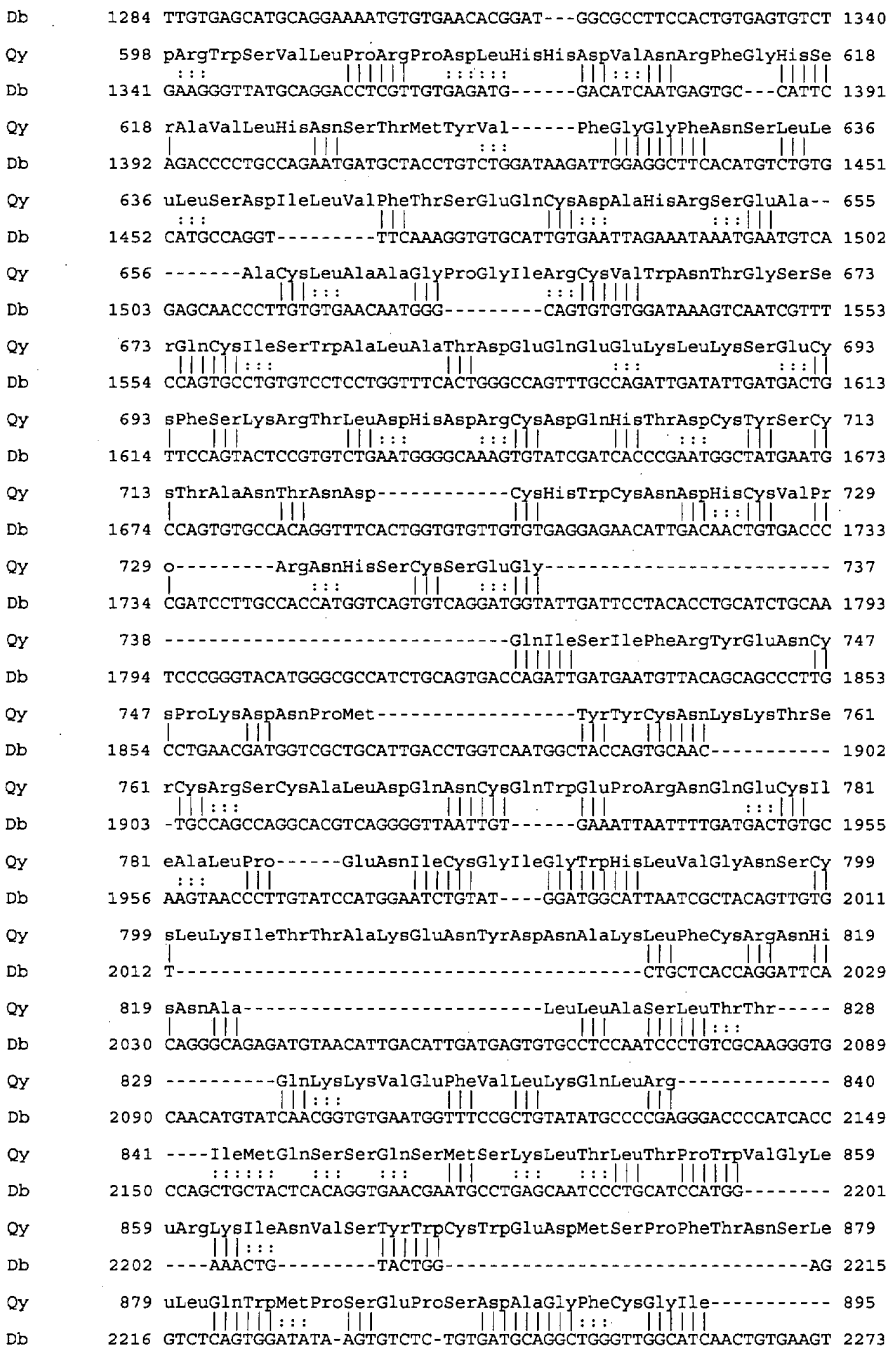
Qy        501 TyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrValHis-------Gl 518
Db       1170 ---------ACTGTGTGACACCAACCCCCTAAATGGGCAATATATTTGCACCTGCCCACA 1220

Qy        518 yGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAs 538
Db       1221 AGGCTACAAAGGGGCTGACTGCACAGAA----------------------------GA 1250

Qy        538 pValAspThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeuHisTh 558
Db       1251 TGTGGATGAATGTGCCATGGCC------------------------------------ 1272

Qy        558 rAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSe 578
Db       1273 ----------------------------------------------AATAGCAATCC 1283

Qy        578 rMetSerHisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAs 598
```





```
Qy       1119 nThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLe 1139
              :::                 :::|||          |||
Db       3276 ATCTGGATGGGCTGGTGCCTATTGTGACGTGCCCAATGTCTCT---------------- 3318

Qy       1139 uCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThr 1158
              |||::::::           :::|||      :::          |||   ::::::
Db       3319 -TGTGACATAGCAGCCTCCAGGAGAGGTGTGCTTGTTGAACACTTGTGCCAGCACTCA 3375
```

Search completed: March 1, 2004, 07:12:55
Job time : 441 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 1, 2004, 05:13:11 ; Search time 11558 Seconds
(without alignments)
5358.816 Million cell updates/sec

Title: US-09-787-097-12
Perfect score: 1429
Sequence: 1 MVAAAAATEARLRRRTAATA.....KSGAVRNRKQQPPAQPGTCI 1429

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Word size: 1

Total number of hits satisfying chosen parameters: 6934743

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09787097/runat_26022004_063014_23748/app_query.fasta_1.1607
-DB=GenEmbl -QFMT=fastap -SUFFIX=oligo.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi -LIST=100
-DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=75 -MODE=LOCAL -OUTFMT=pto
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09787097_@CGN_1_1_7794_@runat_26022004_063014_23748 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1267	88.7	3819	9	AF106861	AF106861 Homo sapi
	2	1141	79.8	8589	6	AR164816	AR164816 Sequence
	3	979	68.5	4072	6	AR164817	AR164817 Sequence
	4	874	61.2	3597	9	AF034957	AF034957 Homo sapi
	5	564	39.5	2578	9	AK000356	AK000356 Homo sapi
	6	503	35.2	6370	6	AR164815	AR164815 Sequence
	7	496	34.7	2625	6	AR164818	AR164818 Sequence
	8	452	31.6	5632	9	AB011120	AB011120 Homo sapi
	9	252	17.6	4493	4	AF531101	AF531101 Bos tauru
	10	218	15.3	4284	10	AB062913	AB062913 Mesocrice
	11	206	14.4	4313	10	AF119821	AF119821 Mus muscu
	12	206	14.4	8739	10	AB038387	AB038387 Rattus no
	13	206	14.4	8827	6	AR164807	AR164807 Sequence
	14	199	13.9	5683	10	AB093245	AB093245 Mus muscu
	15	193	13.5	6083	10	AF116897	AF116897 Mus muscu
	16	181	12.7	543	6	AX778545	AX778545 Sequence
c	17	130	9.1	110000	2	BX004762_0	BX004762 Mus muscu
c	18	130	9.1	110000	2	BX004762_1	Continuation (2 of
	19	130	9.1	110261	10	AL807401	AL807401 Mouse DNA
	20	130	9.1	335799	2	BX005139	BX005139 Mus muscu
	21	104	7.3	315	4	AF301008	AF301008 Bos tauru
	22	94	6.6	2419	6	AR164813	AR164813 Sequence
	23	92	6.4	1012	6	E26747	E26747 Structure a
	24	92	6.4	4500	10	AB038388	AB038388 Rattus no
c	25	82	5.7	246	6	AR273223	AR273223 Sequence
c	26	82	5.7	246	6	AR276804	AR276804 Sequence
c	27	82	5.7	246	6	AR407079	AR407079 Sequence
c	28	82	5.7	246	6	AX368256	AX368256 Sequence
	29	79	5.5	110000	2	AL954766_2	Continuation (3 of
	30	79	5.5	110000	2	BX004762_3	Continuation (4 of
	31	77	5.4	3257	9	H009384S15	AF218903 Homo sapi
	32	77	5.4	104907	9	HSDJ741H3	AL132773 Human DNA
c	33	77	5.4	118540	2	AC013324	AC013324 Homo sapi
c	34	77	5.4	167357	2	AC055771	AC055771 Homo sapi
	35	77	5.4	206647	9	AP002898	AP002898 Homo sapi
	36	77	5.4	207433	6	AX326809	AX326809 Sequence
	37	76	5.3	1051	6	AR164814	AR164814 Sequence
	38	74	5.2	3589	9	H009384S11	AF218899 Homo sapi
c	39	74	5.2	155869	2	AC023420	AC023420 Homo sapi
	40	73	5.1	1207	10	AB06968S15	AB049236 Rattus no
	41	73	5.1	1728	9	H009384S14	AF218902 Homo sapi
	42	73	5.1	90050	6	AR164810	AR164810 Sequence
	43	73	5.1	189283	10	AL833771	AL833771 Mouse DNA
c	44	73	5.1	245589	2	AC094932	AC094932 Rattus no
	45	69	4.8	2018	9	H009384S08	AF218896 Homo sapi
	46	68	4.8	1576	9	H009384S01	AF218889 Homo sapi
	47	68	4.8	51371	9	AL353193	AL353193 Human DNA
c	48	68	4.8	63513	2	AC074061	AC074061 Homo sapi
	49	68	4.8	155869	2	AC023420	AC023420 Homo sapi
	50	64	4.5	3112	9	H009384S13	AF218901 Homo sapi

51	63	4.4	118540	2	AC013324	AC013324 Homo sapi
52	62	4.3	3038	9	H009384S09	AF218897 Homo sapi
53	57	4.0	1032	9	H009384S06	AF218894 Homo sapi
54	53	3.7	1976	9	H009384S21	AF218909 Homo sapi
55	53	3.7	4862	10	AB062914S2	AB062915 Mesocrice
56	53	3.7	167357	2	AC055771	AC055771 Homo sapi
57	53	3.7	206329	2	AC017113	AC017113 Homo sapi
c 58	51	3.6	226	6	AR164811	AR164811 Sequence
59	51	3.6	1050	10	AB06968S09	AB049230 Rattus no
60	48	3.4	553	11	HS741H3S	AL110053 H.sapiens
61	48	3.4	1232	9	H009384S05	AF218893 Homo sapi
62	46	3.2	1674	9	H009384S16	AF218904 Homo sapi
c 63	45	3.1	221	6	AR164812	AR164812 Sequence
64	45	3.1	1880	10	AB06968S16	AB049237 Rattus no
65	44	3.1	1152	9	H009384S04	AF218892 Homo sapi
66	44	3.1	1474	10	AB06968S22	AB049243 Rattus no
67	44	3.1	73666	9	HSDJ581P3	AL109805 Human DNA
68	42	2.9	237	4	AF273610	AF273610 Sus scrof
69	41	2.9	5036	10	AB06968S27	AB049248 Rattus no
70	41	2.9	5790	9	H009384S27	AF218915 Homo sapi
71	41	2.9	7437	10	MMMAHNY2	AF120318 Mus muscu
c 72	41	2.9	206329	2	AC017113	AC017113 Homo sapi
c 73	41	2.9	223026	2	AF415225	AF415225 Mus muscu
74	40	2.8	1267	9	H009384S12	AF218900 Homo sapi
75	39	2.7	1071	10	AB06968S13	AB049234 Rattus no
76	39	2.7	1163	10	AB06968S26	AB049247 Rattus no
77	39	2.7	1173	10	AB062914S5	AB062918 Mesocrice
78	39	2.7	1394	9	H009384S26	AF218914 Homo sapi
79	37	2.6	3058	9	H009384S03	AF218891 Homo sapi
80	36	2.5	1080	10	AB06968S25	AB049246 Rattus no
81	36	2.5	1216	10	AB062914S4	AB062917 Mesocrice
82	36	2.5	1592	9	H009384S25	AF218913 Homo sapi
83	35	2.4	6740	10	AB06968S01	AB069686 Rattus no
84	35	2.4	238009	2	AC103650	AC103650 Mus muscu
85	35	2.4	317248	2	AC133346	AC133346 Rattus no
86	33	2.3	5973	6	AR164809	AR164809 Sequence
87	31	2.2	249	10	AB06968S20	AB049241 Rattus no
88	31	2.2	541	10	AB062914S1	AB062914 Mesocrice
89	31	2.2	921	10	AB06968S12	AB049233 Rattus no
90	31	2.2	1675	9	H009384S17	AF218905 Homo sapi
91	30	2.1	1155	10	AB06968S03	AB049224 Rattus no
92	30	2.1	1204	10	AB06968S14	AB049235 Rattus no
93	30	2.1	1575	9	H009384S07	AF218895 Homo sapi
94	30	2.1	2262	10	AB06968S05	AB049226 Rattus no
c 95	30	2.1	17056	6	AR164808	AR164808 Sequence
96	29	2.0	2704	10	AB06968S02	AB069687 Rattus no
97	29	2.0	3297	10	BC030872	BC030872 Mus muscu
98	29	2.0	3711	9	H009384S02	AF218890 Homo sapi
99	29	2.0	4140	9	AY442317	AY442317 Homo sapi
100	29	2.0	4463	9	AK127277	AK127277 Homo sapi

ALIGNMENTS

RESULT 1
AF106861
LOCUS AF106861 3819 bp mRNA linear PRI 06-JAN-1999
DEFINITION Homo sapiens attractin-2 (ATRN) mRNA, complete cds.
ACCESSION AF106861
VERSION AF106861.1 GI:4093195
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3819)
AUTHORS Duke-Cohan,J.S., Gu,J., Freeman,G.J. and Schlossman,S.F.
TITLE Cloning of cDNA for attractin-2, identical with that of attractin except for a GC-rich 222 bp 5' insertion
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3819)
AUTHORS Duke-Cohan,J.S., Gu,J., Freeman,G.J. and Schlossman,S.F.
TITLE Direct Submission
JOURNAL Submitted (15-NOV-1998) Division of Tumor Immunology, Dana Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, USA
FEATURES Location/Qualifiers
source 1. .3819
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
gene 1. .3819
/gene="ATRN"
CDS 1. .3819
/gene="ATRN"
/note="similar to Homo sapiens attractin in GenBank Accession Number AF034957"
/codon_start=1
/product="attractin-2"
/protein_id="AAD03057.1"
/db_xref="GI:4093196"
/translation="MVAAAAATEARLRRRTAATAALAGRSGGPHWDWDVTRAGRPGLG
AGLRLPRLLSPPLRPRLLLLLLLLPPPLLLLLPCEAEAAAAAAAAVSGSAAAEAKECD
RPCVNGGRCNPGTGQCVCPAGWVGEQCQHCGGRFRLTGSSGFVTDGPGNYKYKTKCTW
LIEGQPNRIMRLRFNHFATECSWDHLYVYDGDSIYAPLVAAFSGLIVPERDGNETVPE
VVATSGYALLHFFSDAAYNLTGFNITYSFDMCPNNCSGRGECKISNSSETVECECSEN
WKGEACDIPHCTDNCGFPHRGICNSSDVRGCSCFSDWQGPGCSVPVPANQSFWTREEY
SNLKLPRASHKAVVNGNIMWVVGGYMFNHSDYNMVLAYDLASREWLPLNRSVNNVVVR
YGHSLALYKDKIYMYGGKIDPTGNVTNELRVFHIHNESWVLLTPKAKEQYAVVGHSAH
IVTLKNGRVVMLVIFGHCPLYGYISNVQEYDLDKNTWSILHTQGALVQGGYGHSSVYD
HRTRALYVHGGYKAFSANKYRLADDLYRYDVDTQMWTILKDSRFFRYLHTAVIVSGTM
LVFGGNTHNDTSMSHGAKCFSSDFMAYDIACDRWSVLPRPDLHHDVNRFGHSAVLHNS
TMYVFGGFNSLLLSDILVFTSEQCDAHRSEAACLAAGPGIRCVWNTGSSQCISWALAT
DEQEEKLKSECFSKRTLDHDRCDQHTDCYSCTANTNDCHWCNDHCVPRNHSCSEGQIS
IFRYENCPKDNPMYYCNKKTSCRSCALDQNCQWEPRNQECIALPENICGIGWHLVGNS
CLKITTAKENYDNAKLFCRNHNALLASLTTQKKVEFVLKQLRIMQSSQSMSKLTLTPW
VGLRKINVSYWCWEDMSPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLN
GSVCERPANHSAKQCRTPCALRTACGDCTSGSSECMWCSNMKQCVDSNAYVASFPFGQ
CMEWYTMSTCPPENCSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKMPSQ
APTGNFYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCENLTTGKH
CETCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGDECQLCEVENRY
QGNPLRGTCYYTLLIDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASKNFN
LNITWAASFSAGTQAGEEMPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTW
PIKIQVQTEQ"
ORIGIN

Alignment Scores:

Pred. No.:	0	Length:	3819
Score:	1267.00	Matches:	1267
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	88.66%	Indels:	0
DB:	9	Gaps:	0

US-09-787-097-12 (1-1429) x AF106861 (1-3819)

```
Qy          1 MetValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 60

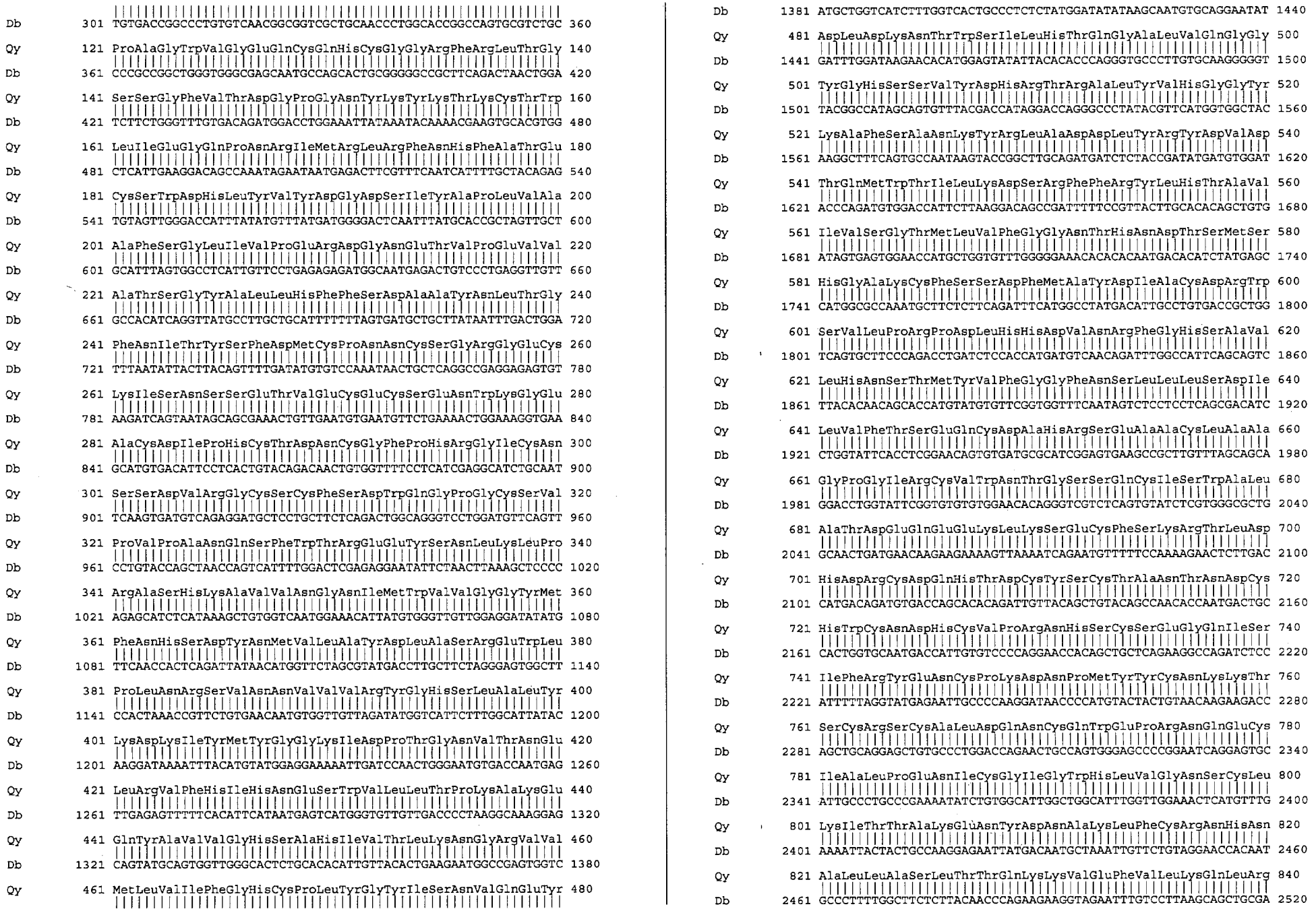
Qy         21 AlaLeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArg 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGGGACTGGGACGTGACCAGGGCTGGGAGG 120

Qy         41 ProGlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSerProProLeuArgProArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGGGGCTGGGGGCCGGGCTGCGCCTCCCGCGGCTGCTGTCTCCACCGCTGCGGCCACGG 180

Qy         61 LeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeuProCysGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGCTGCTGCTGCTGTTGTTGCTCCCGCCGCCGCTGTTGCTGCTGCTGCTGCCCTGTGAG 240

Qy         81 AlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCGAGGCCGCGGCGGCGGCGGCGGCGGTGTCGGGCTCAGCCGCAGCCGAGGCCAAGGAA 300

Qy        101 CysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCys 120
```


Db 301 TGTGACCGGCCCTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGC 360
Qy 121 ProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGly 140
Db 361 CCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGA 420
Qy 141 SerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrp 160
Db 421 TCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGG 480
Qy 161 LeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGlu 180
Db 481 CTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAG 540
Qy 181 CysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAla 200
Db 541 TGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCT 600
Qy 201 AlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGluValVal 220
Db 601 GCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTT 660
Qy 221 AlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGly 240
Db 661 GCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGA 720
Qy 241 PheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCys 260
Db 721 TTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGT 780
Qy 261 LysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsnTrpLysGlyGlu 280
Db 781 AAGATCAGTAATAGCAGCGAAACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAA 840
Qy 281 AlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsn 300
Db 841 GCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAAT 900
Qy 301 SerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerVal 320
Db 901 TCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTT 960
Qy 321 ProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuPro 340
Db 961 CCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCC 1020
Qy 341 ArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMet 360
Db 1021 AGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATG 1080
Qy 361 PheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeu 380
Db 1081 TTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTT 1140
Qy 381 ProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyr 400
Db 1141 CCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATAC 1200
Qy 401 LysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsnValThrAsnGlu 420
Db 1201 AAGGATAAAATTTACATGTATGGAGGAAAAATTGATCCAACTGGGAATGTGACCAATGAG 1260
Qy 421 LeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGlu 440
Db 1261 TTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGGAG 1320
Qy 441 GlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValVal 460
Db 1321 CAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTC 1380
Qy 461 MetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyr 480
Db 1381 ATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATAT 1440
Qy 481 AspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGly 500
Db 1441 GATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGT 1500
Qy 501 TyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyr 520
Db 1501 TACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTAC 1560
Qy 521 LysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAsp 540
Db 1561 AAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGAT 1620
Qy 541 ThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaVal 560
Db 1621 ACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTG 1680
Qy 561 IleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSer 580
Db 1681 ATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGC 1740
Qy 581 HisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrp 600
Db 1741 CATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGG 1800
Qy 601 SerValLeuProArgProAspLeuHisHisAspValAsnArgPheGlyHisSerAlaVal 620
Db 1801 TCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTC 1860
Qy 621 LeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIle 640
Db 1861 TTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATC 1920
Qy 641 LeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAla 660
Db 1921 CTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCA 1980
Qy 661 GlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeu 680
Db 1981 GGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTG 2040
Qy 681 AlaThrAspGluGlnGluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAsp 700
Db 2041 GCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGAC 2100
Qy 701 HisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCys 720
Db 2101 CATGACAGATGTGACCAGCACACAGATTGTTACAGCTGTACAGCCAACACCAATGACTGC 2160
Qy 721 HisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSer 740
Db 2161 CACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCC 2220
Qy 741 IlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThr 760
Db 2221 ATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCCATGTACTACTGTAACAAGAAGACC 2280
Qy 761 SerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCys 780
Db 2281 AGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGC 2340
Qy 781 IleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeu 800
Db 2341 ATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTG 2400
Qy 801 LysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsn 820
Db 2401 AAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAAT 2460
Qy 821 AlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArg 840
Db 2461 GCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGA 2520

```
Qy        841 IleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArg 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 ATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGG 2580

Qy        861 LysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeu 880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 AAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTA 2640

Qy        881 GlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSer 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 CAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGT 2700

Qy        901 ThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArg 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 ACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGG 2760

Qy        921 ProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGly 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 CCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGA 2820

Qy        941 AspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAsp 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 GATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGAC 2880

Qy        961 SerAsnAlaTyrValAlaSerProProHeGlyGlnCysMetGluTrpTyrThrMetSer 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 TCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGC 2940

Qy        981 ThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPro 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 ACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCA 3000

Qy       1001 GlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySer 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 GGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCC 3060

Qy       1021 TyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnPro 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 TATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCC 3120

Qy       1041 LeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCys 1060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 CTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGT 3180

Qy       1061 ProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCys 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 CCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGT 3240

Qy       1081 GluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspPro 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 GAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCC 3300

Qy       1101 ThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThr 1120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 ACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACC 3360

Qy       1121 AsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCys 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 AACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGT 3420

Qy       1141 GluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeu 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 GAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTT 3480

Qy       1161 IleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIle 1180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3481 ATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATC 3540

Qy       1181 AsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSer 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3541 AATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCC 3600

Qy       1201 LysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGly 1220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3601 AAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGA 3660

Qy       1221 GluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsn 1240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3661 GAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAAT 3720

Qy       1241 GluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPhe 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 GAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTC 3780

Qy       1261 ThrTrpProIleLysIleGln 1267
              |||||||||||||||||||||
Db       3781 ACCTGGCCCATCAAAATTCAG 3801
```

RESULT 2
AR164816

LOCUS	AR164816	8589 bp	DNA	linear	PAT 17-OCT-2001

DEFINITION Sequence 14 from patent US 6274339.
ACCESSION AR164816
VERSION AR164816.1 GI:16238025
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 8589)
AUTHORS Moore,K. and Nagle,D.Lynn.
TITLE Methods and compositions for the diagnosis and treatment of body weight disorders, including obesity
JOURNAL Patent: US 6274339-A 14 14-AUG-2001;

FEATURES	Location/Qualifiers
source	1. .8589
	/organism="unknown"
	/mol_type="unassigned DNA"

ORIGIN

Alignment Scores:

Pred. No.:	0	Length:	8589
Score:	1141.00	Matches:	1341
Percent Similarity:	99.70%	Conservative:	0
Best Local Similarity:	99.70%	Mismatches:	2
Query Match:	79.85%	Indels:	4
DB:	6	Gaps:	0

US-09-787-097-12 (1-1429) x AR164816 (1-8589)

```
Qy         87 AlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGluCysAspArgProCysVal 106
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 GCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAAGGAATGTGACCGGCCGTGTC 315

Qy        107 AsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTrpValGly 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        316 AACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGC 375

Qy        127 GluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPheValThr 146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        376 GAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACA 435

Qy        147 AspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeuIleGluGlyGlnPro 166
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        436 GATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCA 495

Qy        167 AsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLeu 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        496 AATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCATTTA 555

Qy        187 TyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaAlaPheSerGlyLeuIle 206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 TATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATT 615

Qy        207 ValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGlyTyrAla 226
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        616 GTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCC 675

Qy        227 LeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThrTyrSer 246

Db        676 TTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTACAGT 735

Qy        247 PheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSerSer 266

Db        736 TTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGC 795

Qy        267 Glu-ThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIleProHi 286

Db        796 GA-TACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCA 854

Qy        286 sCysThrAspAsnCysGlyProHisArgGlyIleCysAsnSerSerAspValArgGl 306

Db        855 CTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGG 914

Qy        306 yCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAlaAsnGl 326

Db        915 ATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCA 974

Qy        326 nSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHisLysAl 346

Db        975 GTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGC 1034

Qy        346 aValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHisSerAspTy 366

Db       1035 TGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTA 1094

Qy        366 rAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArgServa 386

Db       1095 TAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGT 1154

Qy        386 lAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyrMe 406

Db       1155 GAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACAT 1214

Qy        406 tTyrGlyGlyLysIleAspPro-ThrGlyAsnValThrAsnGluLeuArgValPheHisI 426

Db       1215 GTATGGAGGAAAAATTGATTC-AACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACA 1273

Qy        426 leHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaValValG 446

Db       1274 TTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTG 1333

Qy        446 lyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValIlePheG 466

Db       1334 GGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTG 1393

Qy        466 lyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsnT 486

Db       1394 GTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACA 1453

Qy        486 hrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisSerServ 506

Db       1454 CATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTG 1513

Qy        506 alTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaA 526

Db       1514 TTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCCA 1573

Qy        526 snLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetTrpThrI 546

Db       1574 ATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCA 1633

Qy        546 leLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerGlyThrM 566

Db       1634 TTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACCA 1693

Qy        566 etLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysP 586

Db       1694 TGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATGCT 1753

Qy        586 heSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuProArgP 606

Db       1754 TCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGAC 1813

Qy        606 roAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrM 626

Db       1814 CTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACAACAGCACCA 1873

Qy        626 etTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSerG 646

Db       1874 TGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCGG 1933

Qy        646 luGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgC 666

Db       1934 AACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGT 1993

Qy        666 ysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnG 686

Db       1994 GTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAAG 2053

Qy        686 luGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArgCysAspG 706

Db       2054 AAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACC 2113

Qy        706 lnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAspH 726

Db       2114 AGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGCAATGACC 2173

Qy        726 isCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGluA 746

Db       2174 ATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGA 2233

Qy        746 snCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArgSerCysA 766

Db       2234 ATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTG 2293

Qy        766 laLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuProGluA 786

Db       2294 CCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAA 2353

Qy        786 snIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThrThrAlaL 806

Db       2354 ATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCA 2413

Qy        806 ysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSerL 826

Db       2414 AGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTC 2473

Qy        826 euThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGlnSerSerG 846

Db       2474 TTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTC 2533

Qy        846 lnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerT 866

Db       2534 AGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCT 2593

Qy        866 yrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerG 886

Db       2594 ACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTG 2653

Qy        886 luProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysA 906

Db       2654 AGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGG 2713

Qy        906 laAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSerA 926

Db       2714 CTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTG 2773

Qy        926 laLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGlyS 946

Db       2774 CTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGGCA 2833
```

```
Qy       946 erSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrValA 966
Db      2834 GCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGTGG 2893

Qy       966 laSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSerThrCysProProGluA 986
Db      2894 CCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGAAA 2953

Qy       986 snCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCysT 1006
Db      2954 ATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTA 3013

Qy      1006 hrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyProValL 1026
Db      3014 CTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGA 3073

Qy      1026 ysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerSerM 1046
Db      3074 AGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGCA 3133

Qy      1046 etCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnCysA 1066
Db      3134 TGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCA 3193

Qy      1066 snGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrThrG 1086
Db      3194 ACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCACAG 3253

Qy      1086 lyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyGlyLysC 1106
Db      3254 GCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAAT 3313

Qy      1106 ysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysP 1126
Db      3314 GTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGCT 3373

Qy      1126 heCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGluAsnArgT 1146
Db      3374 TCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGAT 3433

Qy      1146 yrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrGlnPheT 1166
Db      3434 ACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTTCA 3493

Qy      1166 hrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIleAsnPheValAlaThrP 1186
Db      3494 CCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTC 3553

Qy      1186 roAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeuA 1206
Db      3554 CTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCTCA 3613

Qy      1206 snIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProValV 1226
Db      3614 ACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTG 3673

Qy      1226 alSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPheAspPheA 1246
Db      3674 TTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTC 3733

Qy      1246 rgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpProIleLysI 1266
Db      3734 GCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAAAA 3793

Qy      1266 leGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePheValThrP 1286
Db      3794 TTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCTGGTACAGTTCTTCGTGACTT 3853

Qy      1286 hePheSerCysPheLeuSerLeuLeuLeuValAlaAlaValValTrpLysIleLysGlnS 1306
Db      3854 TCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGTGGTTTGGAAGATCAAACAAA 3913

Qy      1306 erCysTrpAlaSerArgArgArgGluGlnLeuLeuArgGluMetGlnGlnMetAlaSerA 1326
```

```
Db      3914 GTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGAGATGCAACAGATGGCCAGCC 3973

Qy      1326 rgProPheAlaSerValAsnValAlaLeuGluThrAspGluGluProProAspLeuIleG 1346
Db      3974 GTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTATTG 4033

Qy      1346 lyGlySerIleLysThrValProLysProIleAlaLeuGluProCysPheGlyAsnLysA 1366
Db      4034 GGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGAGCCGTGTTTTGGCAACAAAG 4093

Qy      1366 laAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyGlyIleProProProG 1386
Db      4094 CCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCTGGGTGGCATCCCTCCTCCTG 4153

Qy      1386 lyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetProIleV 1406
Db      4154 GGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGATAG 4213

Qy      1406 alTyrLysGluLysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnProG 1426
Db      4214 TGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGCCTG 4273

Qy      1426 lyThrCysIle 1429
Db      4274 GGACCTGCATC 4284
```

```
RESULT 3
AR164817
LOCUS       AR164817               4072 bp    DNA     linear   PAT 17-OCT-2001
DEFINITION  Sequence 16 from patent US 6274339.
ACCESSION   AR164817
VERSION     AR164817.1  GI:16238027
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 4072)
  AUTHORS   Moore,K. and Nagle,D.Lynn.
  TITLE     Methods and compositions for the diagnosis and treatment of body
            weight disorders, including obesity
  JOURNAL   Patent: US 6274339-A 16 14-AUG-2001;
FEATURES             Location/Qualifiers
     source          1. .4072
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:                0              Length:          4072
Score:                    979.00         Matches:         1179
Percent Similarity:       99.66%         Conservative:    0
Best Local Similarity:    99.66%         Mismatches:      2
Query Match:              68.51%         Indels:          4
DB:                       6              Gaps:            0

US-09-787-097-12 (1-1429) x AR164817 (1-4072)

Qy        87 AlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGluCysAspArgProCysVal 106
Db       256 GCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAAGGAATGTGACCGGCCGTGTGTC 315

Qy       107 AsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTrpValGly 126
Db       316 AACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGC 375

Qy       127 GluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPheValThr 146
Db       376 GAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACA 435

Qy       147 AspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeuIleGluGlyGlnPro 166
Db       436 GATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCA 495
```

```
Qy      167 AsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLeu 186
Db      496 AATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCATTTA 555

Qy      187 TyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaAlaPheSerGlyLeuIle 206
Db      556 TATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATT 615

Qy      207 ValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGlyTyrAla 226
Db      616 GTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCC 675

Qy      227 LeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThrTyrSer 246
Db      676 TTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTACAGT 735

Qy      247 PheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSerSer 266
Db      736 TTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGC 795

Qy      267 Glu-ThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIleProHi 286
Db      796 GA-TACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCA 854

Qy      286 sCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspValArgGl 306
Db      855 CTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGG 914

Qy      306 yCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAlaAsnGl 326
Db      915 ATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCA 974

Qy      326 nSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHisLysAl 346
Db      975 GTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGC 1034

Qy      346 aValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHisSerAspTy 366
Db     1035 TGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTA 1094

Qy      366 rAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSerVa 386
Db     1095 TAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGT 1154

Qy      386 lAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyrMe 406
Db     1155 GAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACAT 1214

Qy      406 tTyrGlyGlyLysIleAspPro-ThrGlyAsnValThrAsnGluLeuArgValPheHisI 426
Db     1215 GTATGGAGGAAAAATTGATTC-AACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACA 1273

Qy      426 leHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaValValG 446
Db     1274 TTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTG 1333

Qy      446 lyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValIlePheG 466
Db     1334 GGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTG 1393

Qy      466 lyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsnT 486
Db     1394 GTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACA 1453

Qy      486 hrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisSerSerV 506
Db     1454 CATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTG 1513

Qy      506 alTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaA 526
Db     1514 TTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCCA 1573

Qy      526 snLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetTrpThrI 546
Db     1574 ATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCA 1633

Qy      546 leLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerGlyThrM 566
Db     1634 TTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACCA 1693

Qy      566 etLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysP 586
Db     1694 TGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATGCT 1753

Qy      586 heSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuProArgP 606
Db     1754 TCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGAC 1813

Qy      606 roAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrM 626
Db     1814 CTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACAACAGCACCA 1873

Qy      626 etTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSerG 646
Db     1874 TGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCGG 1933

Qy      646 luGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgC 666
Db     1934 AACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGT 1993

Qy      666 ysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnG 686
Db     1994 GTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAAG 2053

Qy      686 luGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArgCysAspG 706
Db     2054 AAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACC 2113

Qy      706 lnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAspH 726
Db     2114 AGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGCAATGACC 2173

Qy      726 isCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGluA 746
Db     2174 ATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGA 2233

Qy      746 snCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArgSerCysA 766
Db     2234 ATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTG 2293

Qy      766 laLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuProGluA 786
Db     2294 CCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAA 2353

Qy      786 snIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThrThrAlaL 806
Db     2354 ATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCA 2413

Qy      806 ysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSerL 826
Db     2414 AGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTC 2473

Qy      826 euThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGlnSerSerG 846
Db     2474 TTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTC 2533

Qy      846 lnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerT 866
Db     2534 AGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCT 2593

Qy      866 yrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerG 886
Db     2594 ACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTG 2653

Qy      886 luProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysA 906
```

```
Db       2654 AGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGG 2713

Qy        906 laAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSerA 926

Db       2714 CTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTG 2773

Qy        926 laLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGlyS 946

Db       2774 CTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGGCA 2833

Qy        946 erSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrValA 966

Db       2834 GCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGTGG 2893

Qy        966 laSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSerThrCysProProGluA 986

Db       2894 CCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGAAA 2953

Qy        986 snCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCysT 1006

Db       2954 ATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTA 3013

Qy       1006 hrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyProValL 1026

Db       3014 CTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGA 3073

Qy       1026 ysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerSerM 1046

Db       3074 AGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGCA 3133

Qy       1046 etCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnCysA 1066

Db       3134 TGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCA 3193

Qy       1066 snGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrThrG 1086

Db       3194 ACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCACAG 3253

Qy       1086 lyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyGlyLysC 1106

Db       3254 GCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAAT 3313

Qy       1106 ysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysP 1126

Db       3314 GTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGCT 3373

Qy       1126 heCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGluAsnArgT 1146

Db       3374 TCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGAT 3433

Qy       1146 yrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrGlnPheT 1166

Db       3434 ACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTTCA 3493

Qy       1166 hrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIleAsnPheValAlaThrP 1186

Db       3494 CCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTC 3553

Qy       1186 roAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeuA 1206

Db       3554 CTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCTCA 3613

Qy       1206 snIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProValV 1226

Db       3614 ACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTG 3673

Qy       1226 alSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPheAspPheA 1246

Db       3674 TTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTC 3733

Qy       1246 rgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpProIleLysI 1266

Db       3734 GCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAAAA 3793

Qy       1266 leGln 1267

Db       3794 TTCAG 3798
```

RESULT 4
AF034957

```
LOCUS       AF034957               3597 bp    mRNA    linear   PRI 30-SEP-1998
DEFINITION  Homo sapiens secreted T cell activation protein Attractin
            (attractin) mRNA, complete cds.
ACCESSION   AF034957
VERSION     AF034957.1  GI:3676346
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3597)
  AUTHORS   Duke-Cohan,J.S., Gu,J., McLaughlin,D.F., Xu,Y., Freeman,G.J. and
            Schlossman,S.F.
  TITLE     Attractin (DPPT-L), a member of the CUB family of cell adhesion and
            guidance proteins, is secreted by activated human T lymphocytes and
            modulates immune cell interactions
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 95 (19), 11336-11341 (1998)
  MEDLINE   98409658
   PUBMED   9736737
REFERENCE   2  (bases 1 to 3597)
  AUTHORS   Duke-Cohan,J.S., Gu,J., Ao,Z., McLaughlin,D.F., Freeman,G.J. and
            Schlossman,S.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-NOV-1997) Division of Cancer, Immunology and AIDS,
            Dana Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
            USA
FEATURES             Location/Qualifiers
     source          1. .3597
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     gene            1. .3597
                     /gene="attractin"
     CDS             1. .3597
                     /gene="attractin"
                     /note="secreted T cell activation protein; Method:
                     conceptual translation with partial peptide sequencing;
                     DPPT-L"
                     /codon_start=1
                     /product="Attractin"
                     /protein_id="AAC61902.1"
                     /db_xref="GI:3676347"
                     /translation="MVAAAAATEARLRRRTAATAALAGRSGGPHCVNGGRCNPGTGQC
                     VCPAGWVGEQCQHCGGRFRLTGSSGFVTDGPGNYKYKTKCTWLIEGQPNRIMRLRFNH
                     FATECSWDHLYVYDGDSIYAPLVAAFSGLIVPERDGNETVPEVVATSGYALLHFFSDA
                     AYNLTGFNITYSFDMCPNNCSGRGECKISNSSDTVECECSENWKGEACDIPHCTDNCG
                     FPHRGICNSSDVRGCSCFSDWQGPGCSVPVPANQSFWTREEYSNLKLPRASHKAVVNG
                     NIMWVVGGYMFNHSDYNMVLAYDLASREWLPLNRSVNNVVVRYGHSLALYKDKIYMYG
                     GKIDSTGNVTNELRVFHIHNESWVLLTPKAKEQYAVVGHSAHIVTLKNGRVVMLVIFG
                     HCPLYGYISNVQEYDLDKNTWSILHTQGALVQGGYGHSSVYDHRTRALYVHGGYKAFS
                     ANKYRLADDLYRYDVDTQMWTILKDSRFFRYLHTAVIVSGTMLVFGGNTHNDTSMSHG
                     AKCFSSDFMAYDIACDRWSVLPRPDSTMMSTDLAIPAVLHNSTMYVFGGFNSLLLSDI
                     LVFTSEQCDAHRSEAACLAAGPGIRCVWNTGSSQCISWALATDEQEEKLKSECFSKRT
                     LDHDRCDQHTDCYSCTANTNDCHWCNDHCVPRNHSCSEGQISIFRYENCPKDNPMYYC
                     NKKTSCRSCALDQNCQWEPRNQECIALPENICGIGWHLVGNSCLKITTAKENYDNAKL
                     FCRNHNALLASLTTQKKVEFVLKQLRIMQSSQSMSKLTLTPWVGLRKINVSYWCWEDM
                     SPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLNGSVCERPANHSAKQCR
                     TPCALRTACGDCTSGSSECMWCSNMKQCVDSNAYVASFPFGQCMEWYTMSTCPPENCS
                     GYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKMPSQAPTGNFYPQPLLNSSM
                     CLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCENLTTGKHCETCISGFYGDPTNGG
                     KCQPCKCNGHASLCNTNTGKCFCTTKGVKGDECQLCEVENRYQGNPLRGTCYYTLLID
                     YQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASKNFNLNITWAASFSAGTQAG
                     EEMPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTWPIKIQVQTEQ"
ORIGIN
```

```
Alignment Scores:
Pred. No.:                0              Length:          3597
Score:                    874.00         Matches:         1160
Percent Similarity:       99.49%         Conservative:    0
Best Local Similarity:    99.49%         Mismatches:      3
Query Match:              61.16%         Indels:          6
DB:                       9              Gaps:            0

US-09-787-097-12 (1-1429) x AF034957 (1-3597)

Qy        105 CysValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTrp 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 TGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGG 150

Qy        125 ValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPhe 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 GTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTT 210

Qy        145 ValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeuIleGluGly 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 GTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGA 270

Qy        165 GlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAsp 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 CAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGAC 330

Qy        185 HisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaAlaPheSerGly 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 CATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGC 390

Qy        205 LeuIleValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGly 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 CTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGT 450

Qy        225 TyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 TATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACT 510

Qy        245 TyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsn 264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 TACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAAT 570

Qy        265 SerSerGlu-ThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIl 284
              ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        571 AGCAGCGA-TACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACAT 629

Qy        284 eProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspVa 304
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        630 TCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGT 689

Qy        304 lArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAl 324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        690 CAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGC 749

Qy        324 aAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHi 344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        750 TAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCA 809

Qy        344 sLysAlaValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHisSe 364
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        810 TAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTC 869

Qy        364 rAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnAr 384
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        870 AGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCG 929

Qy        384 gSerValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIl 404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        930 TTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAAT 989

Qy        404 eTyrMetTyrGlyGlyLysIleAspPro-ThrGlyAsnValThrAsnGluLeuArgValP 424
              ||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db        990 TTACATGTATGGAGGAAAAATTGATTC-AACTGGGAATGTGACCAATGAGTTGAGAGTTT 1048

Qy        424 heHisIleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaV 444
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1049 TTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAG 1108

Qy        444 alValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValI 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1109 TGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCA 1168

Qy        464 lePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspL 484
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1169 TCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATA 1228

Qy        484 ysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisS 504
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1229 AGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATA 1288

Qy        504 erSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheS 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1289 GCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCA 1348

Qy        524 erAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetT 544
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1349 GTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGT 1408

Qy        544 rpThrIleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerG 564
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1409 GGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTG 1468

Qy        564 lyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaL 584
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1469 GAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGCCATGGCGCCA 1528

Qy        584 ysCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuP 604
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1529 AATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTC 1588

Qy        604 roArgProAspLeuHisHisAspValAsnArgPheGlyHisSer-AlaValLeuHisAsn 623
              ||||||||||  ||||||||||||||||||||||||||||||| ||||||||||||||||
Db       1589 CCAGACCTGATTC-CACCATGATGTCAACAGATTTGGCCATTCCAGCAGTCTTACACAAC 1647

Qy        624 SerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPhe 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1648 AGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTC 1707

Qy        644 ThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGly 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1708 ACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGT 1767

Qy        664 IleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAsp 683
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1768 ATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGAT 1827

Qy        684 GluGlnGluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArg 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1828 GAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGA 1887

Qy        704 CysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCys 723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1888 TGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGC 1947

Qy        724 AsnAspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArg 743
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1948 AATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGG 2007

Qy        744 TyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArg 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2008 TATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGG 2067

Qy        764 SerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeu 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2068 AGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTG 2127
```

```
Qy        784 ProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThr 803
Db       2128 CCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACT 2187

Qy        804 ThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeu 823
Db       2188 ACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTG 2247

Qy        824 AlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGln 843
Db       2248 GCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAG 2307

Qy        844 SerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsn 863
Db       2308 TCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAAT 2367

Qy        864 ValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMet 883
Db       2368 GTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATG 2427

Qy        884 ProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGly 903
Db       2428 CCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGA 2487

Qy        904 LeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsn 923
Db       2488 CTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAAC 2547

Qy        924 HisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThr 943
Db       2548 CACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACC 2607

Qy        944 SerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAla 963
Db       2608 AGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCC 2667

Qy        964 TyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSerThrCysPro 983
Db       2668 TATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCC 2727

Qy        984 ProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGly 1003
Db       2728 CCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGC 2787

Qy       1004 TrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGly 1023
Db       2788 TGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGA 2847

Qy       1024 ProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsn 1043
Db       2848 CCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAAT 2907

Qy       1044 SerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCys 1063
Db       2908 TCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGC 2967

Qy       1064 GlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeu 1083
Db       2968 CAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTG 3027

Qy       1084 ThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGly 1103
Db       3028 ACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGA 3087

Qy       1104 GlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGly 1123
Db       3088 GGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGC 3147

Qy       1124 LysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGlu 1143
Db       3148 AAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAA 3207

Qy       1144 AsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyr 1163
Db       3208 AATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTAT 3267

Qy       1164 GlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIleAsnPheVal 1183
Db       3268 CAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTG 3327

Qy       1184 AlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPhe 1203
Db       3328 GCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTC 3387

Qy       1204 AsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMet 1223
Db       3388 AACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATG 3447

Qy       1224 ProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPhe 1243
Db       3448 CCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTT 3507

Qy       1244 AspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpPro 1263
Db       3508 GATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCC 3567

Qy       1264 IleLysIleGln 1267
Db       3568 ATCAAAATTCAG 3579
```

```
RESULT 5
AK000356
LOCUS       AK000356               2578 bp    mRNA    linear   PRI 13-SEP-2003
DEFINITION  Homo sapiens cDNA FLJ20349 fis, clone HEP13957, highly similar to
            AB011120 Homo sapiens mRNA for KIAA0548 protein.
ACCESSION   AK000356
VERSION     AK000356.1  GI:7020387
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T.,
            Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Ota,T.,
            Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
            Nakamura,Y., Isogai,T. and Sugano,S.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2578)
  AUTHORS   Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
            Shibahara,T., Tanaka,T. and Nakamura,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science,
            University of Tokyo, Deptment of Virology; Shirokane-dai, 4-6-1,
            Minato-ku, Tokyo 108-8639, Japan (E-mail:flcdna@ims.u-tokyo.ac.jp,
            Tel:81-3-5449-5286, Fax:81-3-5449-5416)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            International Trade and Industry of Japan; cDNA full insert
            sequencing: Research Association for Biotechnology; cDNA library
            construction, 5'- & 3'-end one pass sequencing: Departent of
            Virology and Human Genome Center, Institute of Medical Science,
            University of Tokyo (partly supported by Science and Technology
            Agency).
FEATURES             Location/Qualifiers
     source          1. .2578
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HEP13957"
                     /cell_line="HepG2"
                     /cell_type="hepatoma"
                     /clone_lib="HEP"
                     /note="cloning vector pME18SFL3"
     misc_feature    1. .2578
                     /note="highly similar to AB011120 Homo sapiens mRNA for
```

KIAA0548 protein"
ORIGIN

Alignment Scores:

Pred. No.:	0	Length:	2578
Score:	564.00	Matches:	564
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	39.47%	Indels:	0
DB:	9	Gaps:	0

US-09-787-097-12 (1-1429) x AK000356 (1-2578)

```
Qy        866 TyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSer 885
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 TACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCT 62

Qy        886 GluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLys 905
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 GAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAG 122

Qy        906 AlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSer 925
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGT 182

Qy        926 AlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGly 945
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 GCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGGC 242

Qy        946 SerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrVal 965
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 AGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGTG 302

Qy        966 AlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSerThrCysProProGlu 985
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGAA 362

Qy        986 AsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCys 1005
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 AATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGT 422

Qy       1006 ThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyProVal 1025
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 ACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTG 482

Qy       1026 LysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerSer 1045
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 AAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGC 542

Qy       1046 MetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnCys 1065
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 ATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATGC 602

Qy       1066 AsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrThr 1085
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 AACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCACA 662

Qy       1086 GlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyGlyLys 1105
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 GGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAA 722

Qy       1106 CysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCys 1125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 TGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGC 782

Qy       1126 PheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGluAsnArg 1145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 TTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGA 842

Qy       1146 TyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrGlnPhe 1165
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        843 TACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTTC 902

Qy       1166 ThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIleAsnPheValAlaThr 1185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         903 ACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACT 962

Qy       1186 ProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeu 1205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         963 CCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCTC 1022

Qy       1206 AsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProVal 1225
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1023 AACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTT 1082

Qy       1226 ValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPheAspPhe 1245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1083 GTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTT 1142

Qy       1246 ArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpProIleLys 1265
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1143 CGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAAA 1202

Qy       1266 IleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePheValThr 1285
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1203 ATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCTGGTACAGTTCTTCGTGACT 1262

Qy       1286 PhePheSerCysPheLeuSerLeuLeuLeuValAlaAlaValValTrpLysIleLysGln 1305
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1263 TTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGTGGTTTGGAAGATCAAACAA 1322

Qy       1306 SerCysTrpAlaSerArgArgArgGluGlnLeuLeuArgGluMetGlnGlnMetAlaSer 1325
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1323 AGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGAGATGCAACAGATGGCCAGC 1382

Qy       1326 ArgProPheAlaSerValAsnValAlaLeuGluThrAspGluGluProProAspLeuIle 1345
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1383 CGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTATT 1442

Qy       1346 GlyGlySerIleLysThrValProLysProIleAlaLeuGluProCysPheGlyAsnLys 1365
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1443 GGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGAGCCGTGTTTTGGCAACAAA 1502

Qy       1366 AlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyGlyIleProProPro 1385
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1503 GCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCTGGGTGGCATCCCTCCTCCT 1562

Qy       1386 GlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetProIle 1405
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1563 GGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGATA 1622

Qy       1406 ValTyrLysGluLysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnPro 1425
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1623 GTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGCCT 1682

Qy       1426 GlyThrCysIle 1429
              ||||||||||||
Db        1683 GGGACCTGCATC 1694
```

RESULT 6
AR164815

LOCUS	AR164815	6370 bp	DNA	linear	PAT	17-OCT-2001

DEFINITION Sequence 12 from patent US 6274339.
ACCESSION AR164815
VERSION AR164815.1 GI:16238024
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 6370) .
AUTHORS Moore,K. and Nagle,D.Lynn.
TITLE Methods and compositions for the diagnosis and treatment of body weight disorders, including obesity
JOURNAL Patent: US 6274339-A 12 14-AUG-2001;
FEATURES Location/Qualifiers
source 1. .6370

```
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:                0              Length:          6370
Score:                    503.00         Matches:         516
Percent Similarity:       99.81%         Conservative:    0
Best Local Similarity:    99.81%         Mismatches:      0
Query Match:              35.20%         Indels:          1
DB:                       6              Gaps:            0

US-09-787-097-12 (1-1429) x AR164815 (1-6370)

Qy        914 AsnGlySerValCysGluArgProAlaAsnHisSerAlaLys-GlnCysArgThrProCy 933
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        506 AATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGGCAGTGCCGGACACCATG 565

Qy        933 sAlaLeuArgThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSe 953
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 TGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAG 625

Qy        953 rAsnMetLysGlnCysValAspSerAsnAlaTyrValAlaSerProProheGlyGlnCy 973
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 CAACATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTG 685

Qy        973 sMetGluTrpTyrThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCy 993
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 TATGGAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTG 745

Qy        993 sSerHisCysLeuGluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLy 1013
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 TAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAA 805

Qy       1013 sGlyLysCysIleGluGlySerTyrLysGlyProValLysMetProSerGlnAlaProTh 1033
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 AGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTAC 865

Qy       1033 rGlyAsnPheTyrProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTy 1053
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 AGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATA 925

Qy       1053 rAsnTrpSerPheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAs 1073
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 CAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAA 985

Qy       1073 nGlnSerIleCysGluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIl 1093
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 TCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCAT 1045

Qy       1093 eSerGlyPheTyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGl 1113
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1046 ATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGG 1105

Qy       1113 yHisAlaSerLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLy 1133
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1106 GCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAA 1165

Qy       1133 sGlyAspGluCysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGl 1153
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1166 GGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGG 1225

Qy       1153 yThrCysTyrTyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAs 1173
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1226 AACATGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGA 1285

Qy       1173 pAspArgTyrTyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLe 1193
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1286 TGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTT 1345

Qy       1193 uAspMetPheIleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPh 1213
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1346 GGACATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTT 1405

Qy       1213 eSerAlaGlyThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGl 1233
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1406 CTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGA 1465

Qy       1233 uTyrLysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPh 1253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1466 GTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTT 1525

Qy       1253 ePheValTyrValSerAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHi 1273
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1526 CTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGATTGCCTTCTCTCAGCA 1585

Qy       1273 sSerAsnPheMetAspLeuValGlnPhePheValThrPhePheSerCysPheLeuSerLe 1293
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1586 CAGCAATTTTATGGACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCTTT 1645

Qy       1293 uLeuLeuValAlaAlaValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgAr 1313
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1646 GCTCCTGGTGGCTGCTGTGGTTTGGAAGATCAAACAAAGTTGTTGGGCCTCCAGACGTAG 1705

Qy       1313 gGluGlnLeuLeuArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnVa 1333
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1706 AGAGCAACTTCTTCGAGAGATGCAACAGATGGCCAGCCGTCCCTTTGCCTCTGTAAATGT 1765

Qy       1333 lAlaLeuGluThrAspGluGluProProAspLeuIleGlyGlySerIleLysThrValPr 1353
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1766 CGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCC 1825

Qy       1353 oLysProIleAlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheVa 1373
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1826 CAAACCCATTGCACTGGAGCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGT 1885

Qy       1373 lArgLeuProArgGlyLeuGlyGlyIleProProProGlyGlnSerGlyLeuAlaValAl 1393
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1886 GAGGCTCCCTCGAGGCCTGGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGC 1945

Qy       1393 aSerAlaLeuValAspIleSerGlnGlnMetProIleValTyrLysGluLysSerGlyAl 1413
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1946 CAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGC 2005

Qy       1413 aValArgAsnArgLysGlnGlnProProAlaGlnProGlyThrCysIle 1429
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2006 CGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGCCTGGGACCTGCATC 2054

RESULT 7
AR164818
LOCUS       AR164818               2625 bp    DNA     linear   PAT 17-OCT-2001
DEFINITION  Sequence 18 from patent US 6274339.
ACCESSION   AR164818
VERSION     AR164818.1  GI:16238029
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2625)
  AUTHORS   Moore,K. and Nagle,D.Lynn.
  TITLE     Methods and compositions for the diagnosis and treatment of body
            weight disorders, including obesity
  JOURNAL   Patent: US 6274339-A 18 14-AUG-2001;
FEATURES             Location/Qualifiers
     source          1. .2625
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:                0              Length:          2625
Score:                    496.00         Matches:         696
Percent Similarity:       99.43%         Conservative:    0
Best Local Similarity:    99.43%         Mismatches:      2
Query Match:              34.71%         Indels:          4
DB:                       6              Gaps:            0
```

US-09-787-097-12 (1-1429) x AR164818 (1-2625)

```
Qy        87 AlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGluCysAspArgProCysVal 106
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       256 GCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAAGGAATGTGACCGGCCGTGTGTC 315

Qy       107 AsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTrpValGly 126
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       316 AACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGC 375

Qy       127 GluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPheValThr 146
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       376 GAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACA 435

Qy       147 AspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeuIleGluGlyGlnPro 166
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       436 GATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCA 495

Qy       167 AsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLeu 186
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       496 AATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCATTTA 555

Qy       187 TyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaAlaPheSerGlyLeuIle 206
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       556 TATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATT 615

Qy       207 ValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGlyTyrAla 226
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       616 GTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCC 675

Qy       227 LeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThrTyrSer 246
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       676 TTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTACAGT 735

Qy       247 PheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSerSer 266
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       736 TTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGC 795

Qy       267 Glu-ThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIleProHi 286
             |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       796 GA-TACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCA 854

Qy       286 sCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspValArgGl 306
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       855 CTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGG 914

Qy       306 yCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAlaAsnGl 326
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       915 ATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCA 974

Qy       326 nSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHisLysAl 346
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       975 GTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGC 1034

Qy       346 aValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHisSerAspTy 366
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1035 TGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTA 1094

Qy       366 rAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArgServa 386
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1095 TAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGT 1154

Qy       386 lAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyrMe 406
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1155 GAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACAT 1214

Qy       406 tTyrGlyGlyLysIleAspPro-ThrGlyAsnValThrAsnGluLeuArgValPheHisI 426
             ||||||||||||||||||    ||||||||||||||||||||||||||||||||||||||
Db      1215 GTATGGAGGAAAAATTGATTC-AACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACA 1273

Qy       426 leHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaValValG 446
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1274 TTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTG 1333

Qy       446 lyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValIlePheG 466
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1334 GGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTG 1393

Qy       466 lyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsnT 486
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1394 GTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACA 1453

Qy       486 hrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisSerServ 506
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1454 CATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTG 1513

Qy       506 alTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaA 526
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1514 TTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCCA 1573

Qy       526 snLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetTrpThrI 546
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1574 ATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCA 1633

Qy       546 leLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerGlyThrM 566
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1634 TTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACCA 1693

Qy       566 etLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysP 586
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1694 TGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATGCT 1753

Qy       586 heSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuProArgP 606
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1754 TCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGAC 1813

Qy       606 roAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrM 626
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1814 CTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACAACAGCACCA 1873

Qy       626 etTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSerG 646
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1874 TGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCGG 1933

Qy       646 luGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgC 666
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1934 AACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGT 1993

Qy       666 ysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnG 686
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1994 GTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAAG 2053

Qy       686 luGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArgCysAspG 706
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2054 AAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACC 2113

Qy       706 lnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAspH 726
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2114 AGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGCAATGACC 2173

Qy       726 isCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGluA 746
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2174 ATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGA 2233

Qy       746 snCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArgSerCysA 766
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2234 ATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTG 2293

Qy       766 laLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuPro 784
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2294 CCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCC 2349
```

RESULT 8
AB011120
LOCUS AB011120 5632 bp mRNA linear PRI 10-APR-1998

```
DEFINITION  Homo sapiens mRNA for KIAA0548 protein, partial cds.
ACCESSION   AB011120
VERSION     AB011120.1  GI:3043619
KEYWORDS    KIAA0548 protein.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Nagase,T., Ishikawa,K., Miyajima,N., Tanaka,A., Kotani,H.,
            Nomura,N. and Ohara,O.
  TITLE     Prediction of the coding sequences of unidentified human genes. IX.
            The complete sequences of 100 new cDNA clones from brain which can
            code for large proteins in vitro
  JOURNAL   DNA Res. 5 (1), 31-39 (1998)
  MEDLINE   98290545
   PUBMED   9628581
REFERENCE   2  (bases 1 to 5632)
  AUTHORS   Ohara,O., Nagase,T. and Ishikawa,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-FEB-1998) Osamu Ohara, Kazusa DNA Research Institute,
            DNA Technology; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            (E-mail:cdnainfo@kazusa.or.jp, Tel:+81-438-52-3913,
            Fax:+81-438-52-3914)
FEATURES             Location/Qualifiers
     source          1. .5632
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HH0769"
                     /sex="male"
                     /tissue_type="brain"
                     /clone_lib="pBluescriptII SK plus"
     gene            1. .5632
                     /gene="KIAA0548"
     CDS             <1. .1361
                     /gene="KIAA0548"
                     /codon_start=3
                     /product="KIAA0548 protein"
                     /protein_id="BAA25474.1"
                     /db_xref="GI:3043620"
                     /translation="TMSTCPPENCSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSY
                     KGPVKMPSQAPTGNFYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSICEK
                     CENLTTGKHCETCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGDEC
                     QLCEVENRYQGNPLRGTCYYTLLIDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDM
                     FINASKNFNLNITWAASFSAGTQAGEEMPVVSKTNIKEYKDSFSNEKFDFRNHPNITF
                     FVYVSNFTWPIKIQIAFSQHSNFMDLVQFFVTFFSCFLSLLLVAAVVWKIKQSCWASR
                     RREQLLREMQQMASRPFASVNVALETDEEPPDLIGGSIKTVPKPIALEPCFGNKAAVL
                     SVFVRLPRGLGGIPPPGQSGLAVASALVDISQQMPIVYKEKSGAVRNRKQQPPAQPGT
                     CI"
ORIGIN

Alignment Scores:
Pred. No.:                0              Length:          5632
Score:                    452.00         Matches:         452
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              31.63%         Indels:          0
DB:                       9              Gaps:            0

US-09-787-097-12 (1-1429) x AB011120 (1-5632)

Qy        978 ThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeu 997
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 ACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTG 62

Qy        998 GluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIle 1017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 GAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATA 122

Qy       1018 GluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyr 1037
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTAT 182

Qy       1038 ProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPhe 1057
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 CCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 242

Qy       1058 IleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCys 1077
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 ATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGT 302

Qy       1078 GluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyr 1097
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTAC 362

Qy       1098 GlyAspProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeu 1117
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTG 422

Qy       1118 CysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCys 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 TGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGC 482

Qy       1138 GlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyr 1157
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 CAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTAT 542

Qy       1158 ThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyr 1177
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 ACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTAC 602

Qy       1178 ThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIle 1197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 ACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATC 662

Qy       1198 AsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThr 1217
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 AATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACC 722

Qy       1218 GlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSer 1237
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 CAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGT 782

Qy       1238 PheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrVal 1257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 TTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTC 842

Qy       1258 SerAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMet 1277
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        843 AGTAATTTCACCTGGCCCATCAAAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATG 902

Qy       1278 AspLeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuLeuValAla 1297
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        903 GACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCT 962

Qy       1298 AlaValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeuLeu 1317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        963 GCTGTGGTTTGGAAGATCAAACAAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTT 1022

Qy       1318 ArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThr 1337
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1023 CGAGAGATGCAACAGATGGCCAGCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACA 1082

Qy       1338 AspGluGluProProAspLeuIleGlyGlySerIleLysThrValProLysProIleAla 1357
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1083 GATGAGGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCA 1142

Qy       1358 LeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProArg 1377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1143 CTGGAGCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGA 1202

Qy       1378 GlyLeuGlyGlyIleProProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuVal 1397
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1203 GGCCTGGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTG 1262
```

```
Qy       1398 AspIleSerGlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAsnArg 1417
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1263 GACATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGG 1322

Qy       1418 LysGlnGlnProProAlaGlnProGlyThrCysIle 1429
             |||||||||||||||||||||||||||||||||||||
Db       1323 AAGCAGCAGCCCCCTGCACAGCCTGGGACCTGCATC 1358

RESULT 9
AF531101
LOCUS       AF531101                4493 bp    mRNA    linear   MAM 19-AUG-2002
DEFINITION  Bos taurus attractin mRNA, complete cds.
ACCESSION   AF531101
VERSION     AF531101.1  GI:22297307
KEYWORDS    .
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 4493)
  AUTHORS   Graphodatskaya,D., Joerg,H. and Stranzinger,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUL-2002) Animal Sciences, ETH, Tannenstrasse 1,
            Zurich, CH 8092, Switzerland
FEATURES             Location/Qualifiers
     source          1. .4493
                     /organism="Bos taurus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9913"
                     /chromosome="13"
     CDS             52. .4299
                     /note="transmembrane form; mahogany protein"
                     /codon_start=1
                     /product="attractin"
                     /protein_id="AAM95445.1"
                     /db_xref="GI:22297308"
                     /translation="MVAAATEARLRRTVVTPAPTSRNGRRCRDWDATRAGRPGLRTGL
                     RLPRRWRRLLLLLLPPLLLLPWAAEAAAAAASVSGSAAAEAKECDRPCVNGGRCNPGT
                     GQCVCPAGWVGEQCQHCGGRFRLTGSSGFITDGPGNYKYKTKCTWLIEGQPNRIMRLR
                     FNHFATECSWDHLYVYDGDSIYAPLVAAFSGLIVPERDGNETVPEVVATSGYALLHFF
                     SDAAYNLTGFNITYNFDMCPNNCSGRGECKISNSSNTVQCECSENWKGEACDIPHCVN
                     NCGFPHRGICNSSDVRGCSCFSEWQGPGCSVPVPANQSFWTREEYSDLKLPRASHKAV
                     VNGNIMWVVGGYMFNHSDYNMVLAYDLASREWLALNRSVNSVVVRYGHSLALYQDKIY
                     MYGGKIDSTGNVTNELRVFHIHNESWVLLSPKAKEQYAVVGHSAHIVTLKNGRVVMLV
                     IFGHCPLYGYISSVQEYDLDKNTWSILHTQGALVQGGYGHSSVYDDRTKALYIHGGYK
                     AFSANKYRLADDLYRYDVDTQMWTILKDSRFFRYLHTAVIVSGTMLVFGGNTHNDTSM
                     SHGAKCFSSDFMAYDIDCDRWSVLPRPDLHHDVNRFGHSAVLYNRTMYVFGGFNSLLL
                     SDILVFTSEQCEAHQSEAACLAAGPGVRCMWDTGSSQCVSWELAPEAQEKIKSECFSK
                     RIFDHDRCDQLTDCYSCTANTNGCQWCNDHCVPTNHSCTEGQISIFKYDHCPKDNPMY
                     YCNKKTSCRSCALDQNCQWEPRNQECIALPENICGIGWHLVGNSCLKITTARETYDNA
                     KLSCRNHNAFLASLTTQKKVEFVLKQLRIMQSSQSMSKFTLTPWVGLRKINVSYWCWE
                     DMSPFTNSSLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLNGSVCERPANHSAKQ
                     CRTPCALRTACGECTSGSSECMWCSNMKQCVDSNAYVASFPFGQCMEWYTMSSCPPEN
                     CSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKMPSQGPTGNSYPQPLLNS
                     SMCLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCENLTTGKHCETCISGFYGDPTN
                     GGKCQPCRCNGHASLCNTNTGKCFCTTKGVKGDECQLCEVENRYQGNPLKGTCYYTLL
                     IDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASKNFNLNITWAASFSAGTQ
                     AGEEMPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTWPIKIQIAFSQHSNF
                     MDLVQFFVTFFSCFLSLLLVAAVVWKIKQSCWASRRREQLLREMQQMASRPFASVNVA
                     LETDEEPPDLIGGSIKTVPKPIALEPCFGNKAAVLSVFVRLPRGLGGIPPPGQSGLAV
                     ASALVDISQQMPVVYKEKSGAVRNRKQQPPAQPGTCI"
ORIGIN

Alignment Scores:
Pred. No.:              9.76e-232       Length:         4493
Score:                  252.00          Matches:        574
Percent Similarity:     97.29%          Conservative:   0
Best Local Similarity:  97.29%          Mismatches:     8
Query Match:            17.63%          Indels:         16
DB:                     4               Gaps:           0

US-09-787-097-12 (1-1429) x AF531101 (1-4493)
```

```
Qy        823 LeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMet 842
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2476 TTGGCTTCTCTTACAACTCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATCATG 2535

Qy        843 GlnSerSerGlnSerMetSerLysLeu-ThrLeuThrProTrpValGlyLeuArgLysIl 862
              |||||||||||||||||||||||||    ||||||||||||||||||||||||||||||||
Db       2536 CAGTCATCACAGAGCATGTCCAAGTT-CACCTTAACCCCATGGGTTGGCCTTCGGAAGAT 2594

Qy        862 eAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeu-LeuGlnT 882
              |||||||||||||||||||||||||||||||||||||||||||||||||||    |||||||
Db       2595 CAATGTGTCCTATTGGTGCTGGGAAGATATGTCCCCGTTCACAAATAGTTC-ACTACAGT 2653

Qy        882 rpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrA 902
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2654 GGATGCCATCTGAGCCCAGTGATGCTGGATTCTGTGGGATCTTGTCAGAACCCAGTACTC 2713

Qy        902 rgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProA 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2714 GGGGCTTAAAGGCTGCAACCTGCATCAATCCACTCAATGGTAGTGTCTGTGAAAGGCCTG 2773

Qy        922 laAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCys-GlyAsp 941
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db       2774 CTAACCACAGTGCCAAGCAGTGTCGGACCCCTTGTGCCTTGCGGACAGCCTGTGGGGAG- 2832

Qy        942 CysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSer 961
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2833 TGTACCAGTGGCAGCTCGGAGTGCATGTGGTGCAGCAACATGAAGCAGTGCGTGGACTCC 2892

Qy        962 AsnAlaTyrValAlaSerProProGlyGlnCysMetGluTrpTyrThrMetSerThr 981
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2893 AATGCCTACGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAGTGGTATACCATGAGCAG- 2951

Qy        982 -CysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGl 1001
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2952 CTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCACTGCTTGGAGCAACCAGG 3011

Qy       1001 yCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTy 1021
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3012 CTGTGGCTGGTGCACTGATCCTAGCAATACGGGCAAAGGGAAATGCATAGAGGGCTCCTA 3071

Qy       1021 rLysGlyProValLysMetProSerGlnAla-ProThrGlyAsnPhe-TyrProGlnPro 1040
              |||||||||||||||||||||||||||||    |||||||||    ||| ||||||||||||
Db       3072 TAAAGGACCAGTGAAGATGCCTTCTCAGGG-CCCTACAGGAAA-TTCCTACCCACAGCCT 3129

Qy       1041 LeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCys 1060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3130 CTTCTCAATTCCAGCATGTGTCTGGAGGACAGCAGATATAACTGGTCTTTCATTCATTGT 3189

Qy       1061 ProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCys 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3190 CCAGCTTGCCAGTGCAATGGCCACAGCAAATGCATTAATCAGAGTATCTGTGAGAAGTGT 3249

Qy       1081 GluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspPro 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3250 GAGAACCTGACCACGGGCAAGCACTGTGAGACCTGCATATCTGGCTTCTATGGTGATCCA 3309

Qy       1101 ThrAsnGlyGlyLysCysGlnProCysLys-CysAsnGlyHisAlaSerLeuCysAsnTh 1120
              |||||||||||||||||||||||||||||    ||||||||||||||||||||||||||||||
Db       3310 ACCAACGGGGGAAAATGTCAACCATGCAG-GTGCAATGGACATGCGTCGCTATGCAATAC 3368

Qy       1120 rAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCy 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3369 CAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGATGAATGTCAGCTATG 3428

Qy       1140 sGluValGluAsnArgTyrGlnGlyAsnProLeuArg-GlyThrCysTyrTyrThrLeuL 1160
              |||||||||||||||||||||||||||||||||    ||||||||||||||||||||||||
Db       3429 TGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAA-AGGAACGTGTTATTATACTCTTC 3487

Qy       1160 euIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaI 1180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3488 TTATTGACTATCAGTTCACCTTTAGCCTATCCCAGGAAGATGACCGCTATTACACAGCCA 3547
```

```
Qy       1180 leAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaS 1200
Db       3548 TCAATTTTGTGGCGACACCTGATGAACAAAACAGGGATTTGGACATGTTCATCAATGCCT 3607

Qy       1200 erLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaG 1220
Db       3608 CCAAAAACTTCAACCTCAACATCACCTGGGCTGCCAGCTTCTCAGCTGGAACCCAAGCTG 3667

Qy       1220 lyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerA 1240
Db       3668 GAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTA 3727

Qy       1240 snGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnP 1260
Db       3728 ATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATT 3787

Qy       1260 heThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuV 1280
Db       3788 TCACCTGGCCCATCAAAATTCAGATTGCATTCTCCCAGCACAGCAATTTTATGGACCTAG 3847

Qy       1280 alGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuLeuValAlaAlaValV 1300
Db       3848 TACAGTTCTTCGTGACCTTCTTCAGTTGTTTCCTCTCTCTGCTCCTGGTGGCTGCTGTGG 3907

Qy       1300 alTrpLysIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeuLeuArgGluM 1320
Db       3908 TTTGGAAGATCAAACAAAGTTGTTGGGCTTCCAGACGTAGAGAGCAACTTCTTCGAGAGA 3967

Qy       1320 etGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThrAspGluG 1340
Db       3968 TGCAGCAGATGGCCAGTCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGAAG 4027

Qy       1340 luProProAspLeuIleGlyGlySerIleLysThrValProLysProIleAlaLeuGluP 1360
Db       4028 AGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAGCCCATTGCCCTGGAGC 4087

Qy       1360 roCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProArgGlyLeuG 1380
Db       4088 CGTGTTTTGGCAACAAAGCTGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCTCG 4147

Qy       1380 lyGlyIleProProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleS 1400
Db       4148 GTGGAATCCCTCCTCCTGGGCAGTCCGGGCTCGCAGTGGCCAGTGCCCTGGTGGACATTT 4207

Qy       1400 erGlnGlnMetPro 1404
Db       4208 CTCAGCAGATGCCA 4221
```

```
RESULT 10
AB062913
LOCUS       AB062913                4284 bp    mRNA    linear   ROD 16-JAN-2002
DEFINITION  Mesocricetus auratus Atrn mRNA for attractin, complete cds.
ACCESSION   AB062913
VERSION     AB062913.1  GI:16930100
KEYWORDS    .
SOURCE      Mesocricetus auratus (golden hamster)
  ORGANISM  Mesocricetus auratus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
            Mesocricetus.
REFERENCE   1
  AUTHORS   Kuramoto,T., Nomoto,T., Fujiwara,A., Mizutani,M., Sugimura,T. and
            Ushijima,T.
  TITLE     Insertional mutation of the Attractin gene in the black tremor
            hamster
  JOURNAL   Mamm. Genome 13 (1), 36-40 (2002)
  MEDLINE   21635551
   PUBMED   11773967
REFERENCE   2  (bases 1 to 4284)
  AUTHORS   Kuramoto,T. and Ushijima,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-JUN-2001) Takashi Kuramoto, National Cancer Center
            Research Institute, Carcinogenesis Division; Tsukiji 5-1-1,
            Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramot@gan2.ncc.go.jp,
            URL:http://www.ncc.go.jp/research/rat-genome/, Tel:81-3-3542-2511,
            Fax:81-3-5565-1753)
FEATURES             Location/Qualifiers
     source          1. .4284
                     /organism="Mesocricetus auratus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10036"
     gene            1. .4284
                     /gene="Atrn"
     CDS             1. .4284
                     /gene="Atrn"
                     /codon_start=1
                     /product="attractin"
                     /protein_id="BAB72012.1"
                     /db_xref="GI:16930101"
                     /translation="MVAAAAATEALLRGSTAATAAPAGRKVEQYRLCAATGAWRPGPR
                     AGLCLPWVVSRALPPPLLLLLLLFSLLLLPLPREAEAAAVSAAVSGSAAAEAKECDRP
                     CVNGGRCNPGTGQCVCPAGWVGEQCQHCGGRFRLTGSSGFVTDGPGNYKYKTKCTWLI
                     EGQPNRIMRLRFNHFATECSWDHLYVYDGDSIYAPLIAAFSGLIVPERDGNETAPEVT
                     VTSGYALLHFFSDAAYNLTGFNITYNFDMCPNNCSGRGECKSSNSSNTVECECSENWK
                     GESCDIPHCTDNCGFPHRGICNASDVRGCSCSPHWQGPGCSIPVPANQSFWIREEYSD
                     FKLPRASHKAVVNGNIMWVVGGYMFNHSDYSMVLAYDLASREWLPLNHSVNNVVVRYG
                     HSLALHKDKIYMYGGKIDSTGNVTNELRVFHIHNESWVLLTPKAKEQYAVVGHSAHIV
                     TLATGRVVMLVIFGHCPLYGYISIVQEYDLEKNTWSILQTQGALVQGGYGHSSVYDHR
                     TKALYVHGGYKAFSANKYRLADDLYRYDVDTQMWTILKDSRFFRYLHTAVIVSGTMLV
                     FGGNTHNDTSMSHGAKCFSSDFMAYDIACDRWSVLPRPDLHHDVNRFGHSAVLHNSTM
                     YVFGGFNSLLLSDVLVFTSEQCDAHHSEAACVAAGPGIRCLWDTVSSRCTSWELATEE
                     QAEKLKSECFSKRTLDHDRCDQHTDCYSCTANTNDCHWCNDHCVPVNHSCTEGQISIS
                     KYDNCPKDNPMYYCNKKTSCRSCALDQNCQWEPRNQECIALPENICGIGWHLVGNSCL
                     KITTAKENYDNAKLSCRNHNAFLASLTSQKKVEFVLKHLRLMQSSQSMSKLTLTPWVG
                     LRKINVSYWCWEDMSPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLNGS
                     VCERPANHSAKQCRTPCALRTACSECTSSSSECMWCSNMKQCVDSNAYVASFPFGQCM
                     EWYTMSSCPPENCSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPIRMPSQAS
                     TGNVYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCENLTTGKHCE
                     TCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGDECQLCEVENRYQG
                     NPLKGTCYYTLLIDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASKNFNLN
                     ITWATSFSAGTQAGEEMPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTWPI
                     KIQIAFSQHSNFMDLVQFFVTFFSCFLSLLLVAAVVWKIKQSCWASRRREQLLREMQQ
                     MASRPFASVNVALETDEEPPDLIGGSIKTVPKPIALEPCFGNKAAVLSVFVRLPRGLG
                     GIPPPGQSGLAVASALVDISQQMPIVYKEKSGAVRNRKQQPPAQPGTCI"
ORIGIN

Alignment Scores:
Pred. No.:                5.07e-199       Length:          4284
Score:                    218.00          Matches:         579
Percent Similarity:       96.98%          Conservative:    0
Best Local Similarity:    96.98%          Mismatches:      9
Query Match:              15.26%          Indels:          18
DB:                       10              Gaps:            0

US-09-787-097-12 (1-1429) x AB062913 (1-4284)

Qy        842 MetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLys 861
Db       2518 ATGCAGTCATCTCAGAGCATGTCCAAGCTCACTCTGACTCCATGGGTTGGTCTTCGGAAG 2577

Qy        862 IleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGln 881
Db       2578 ATCAATGTGTCTTATTGGTGCTGGGAGGATATGTCTCCATTTACAAATAGTTTGCTGCAG 2637

Qy        882 TrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThr 901
Db       2638 TGGATGCCATCCGAGCCCAGTGATGCTGGTTTCTGTGGGATCTTGTCAGAGCCTAGTACT 2697

Qy        902 ArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro 921
Db       2698 CGGGGATTGAAGGCTGCAACCTGCATCAACCCACTCAATGGCAGTGTTTGTGAAAGGCCT 2757

Qy        922 AlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAsp 941
Db       2758 GCAAACCACAGTGCCAAGCAGTGCCGGACACCGTGTGCCCTGCGGACAGCGTGTAG-TG- 2815
```

```
Qy        942 --CysThrSerGly-SerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAsp 960
              |||||||||    ||||||||||||||||||||||||||||||||||||||||||||||
Db       2816 AGTGCACTAGCAG-CAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTAGAC 2874

Qy        961 SerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSer 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2875 TCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAGTGGTATACAATGAGC 2934

Qy        981 Thr-CysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPr 1000
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2935 AG-CTGCCCACCTGAAAATTGCTCTGGCTACTGTACCTGCAGCCATTGCTTGGAGCAGCC 2993

Qy       1000 oGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySe 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2994 AGGCTGTGGCTGGTGTACTGATCCTAGCAATACTGGGAAAGGGAAATGTATCGAGGGCAG 3053

Qy       1020 rTyrLysGlyProValLys-MetProSerGlnAlaPro-ThrGlyAsnPhe-TyrProGl 1039
              |||||||||||||    || ||||||||||||||    |||||||||    ||||||||
Db       3054 CTATAAAGGACCCAT-AAGGATGCCTTCACAGGCTTC-TACAGGAAATGT-GTATCCACA 3110

Qy       1039 nProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHi 1059
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3111 GCCTCTTCTGAACTCTAGCATGTGTTTAGAGGACAGCAGATACAACTGGTCTTTCATTCA 3170

Qy       1059 sCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLy 1079
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3171 CTGTCCAGCTTGCCAGTGCAATGGACACAGCAAATGTATTAACCAGAGTATCTGTGAGAA 3230

Qy       1079 sCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAs 1099
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3231 GTGTGAGAACCTGACCACTGGCAAGCACTGTGAGACCTGCATATCTGGCTTCTATGGTGA 3290

Qy       1099 pProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAs 1119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3291 TCCGACTAATGGAGGCAAATGTCAGCCATGTAAGTGCAACGGGCATGCATCGCTGTGCAA 3350

Qy       1119 nThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLe 1139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3351 TACCAACACCGGCAAGTGCTTCTGCACCACCAAGGGTGTCAAGGGAGATGAGTGCCAGCT 3410

Qy       1139 uCysGluValGluAsnArgTyrGlnGlyAsnProLeuArg-GlyThrCysTyrTyrThrL 1159
              |||||||||||||||||||||||||||||||||||||    ||||||||||||||||||||
Db       3411 ATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAA-AGGAACATGTTATTATACCC 3469

Qy       1159 euLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrA 1179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3470 TTCTCATTGACTATCAGTTCACTTTTAGCCTGTCCCAGGAAGACGATCGCTATTATACAG 3529

Qy       1179 laIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnA 1199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3530 CCATCAATTTTGTGGCCACTCCTGATGAACAAAACAGGGATTTGGACATGTTTATCAATG 3589

Qy       1199 laSerLysAsnPheAsnLeuAsnIleThrTrpAlaAla-SerPheSerAlaGlyThrGln 1218
              ||||||||||||||||||||||||||||||||||    ||||||||||||||||||||
Db       3590 CCTCAAAAAACTTCAACCTCAACATCACCTGGGCTAC-CAGCTTCTCAGCTGGAACCCAG 3648

Qy       1219 AlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSerPhe 1238
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3649 GCAGGAGAAGAAATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGCTTC 3708

Qy       1239 SerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrValSer 1258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3709 TCTAATGAGAAATTTGATTTTCGCAACCATCCAAACATTACTTTCTTTGTTTATGTCAGT 3768

Qy       1259 AsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAsp 1278
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3769 AATTTCACCTGGCCCATCAAAATTCAGATTGCCTTCTCCCAGCACAGCAATTTTATGGAC 3828

Qy       1279 LeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuLeuValAlaAla 1298
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3829 CTGGTACAGTTCTTCGTGACTTTCTTCAGCTGTTTCCTCTCTCTGCTCCTAGTGGCTGCA 3888

Qy       1299 ValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeuLeuArg 1318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3889 GTGGTTTGGAAGATCAAGCAGAGCTGCTGGGCATCCAGGAGGAGGGAGCAACTTCTTCGG 3948

Qy       1319 GluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThrAsp 1338
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3949 GAGATGCAGCAGATGGCCAGCCGTCCCTTTGCTTCTGTAAATGTTGCCTTGGAAACAGAT 4008

Qy       1339 GluGluProProAspLeuIleGlyGlySerIleLysThrValProLysProIleAlaLeu 1358
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4009 GAAGAGCCTCCTGATCTTATTGGGGGAAGTATAAAGACTGTCCCCAAGCCCATTGCCCTG 4068

Qy       1359 GluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProArgGly 1378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4069 GAGCCCTGCTTTGGTAACAAAGCTGCAGTCCTCTCTGTATTTGTGAGGCTCCCTCGAGGA 4128

Qy       1379 LeuGlyGlyIleProProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAsp 1398
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4129 CTGGGTGGAATCCCTCCTCCTGGTCAGTCAGGTCTTGCTGTGGCCAGTGCCCTGGTGGAC 4188

Qy       1399 IleSerGlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAsnArgLys 1418
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4189 ATTTCTCAACAGATGCCGATAGTATACAAAGAGAAGTCAGGAGCGGTAAGAAACCGGAAG 4248

Qy       1419 GlnGlnProProAlaGlnProGlyThrCysIle 1429
              |||||||||||||||||||||||||||||||||
Db       4249 CAGCAGCCCCCTGCACAGCCTGGAACCTGCATT 4281
```

```
RESULT 11
AF119821
LOCUS       AF119821                4313 bp    mRNA    linear   ROD 16-APR-1999
DEFINITION  Mus musculus attractin (Mgca) mRNA, complete cds.
ACCESSION   AF119821
VERSION     AF119821.1  GI:4585306
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 4313)
  AUTHORS   Gunn,T.M., Miller,K.A., He,L., Hyman,R.W., Davis,R.W., Azarani,A.,
            Schlossman,S.F., Duke-Cohan,J.S. and Barsh,G.S.
  TITLE     The mouse mahogany locus encodes a transmembrane form of human
            attractin
  JOURNAL   Nature 398 (6723), 152-156 (1999)
  MEDLINE   99184160
   PUBMED   10086356
REFERENCE   2  (bases 1 to 4313)
  AUTHORS   Gunn,T.M., Miller,K.A., He,L., Hyman,R.W., Davis,R.W., Azarani,A.,
            Schlossman,S.F., Duke-Cohan,J.S. and Barsh,G.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JAN-1999) HHMI, Stanford, Beckman Center B271A,
            Stanford, CA 94305-5323, USA
FEATURES             Location/Qualifiers
     source          1. .4313
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /chromosome="2"
     gene            1. .4313
                     /gene="Mgca"
                     /note="mahogany candidate"
     CDS             27. .4313
                     /gene="Mgca"
                     /note="mahogany"
                     /codon_start=1
                     /product="attractin"
                     /protein_id="AAD25372.1"
                     /db_xref="GI:4585307"
                     /translation="MVAVAVAASTEARLRGSTTATAAPAGRKGRQHRPCTATGAWRPG
                     PRARLCLPRVLSRALPPPPLLPLLFSLLLLPLPREAEAAAAVAAAVSGSAAAEAKECDR
                     PCVNGGRCNPGTGQCVCPTGWVGEQCQHCGGRFRLTGSSGFVTDGPGNYKYKTKCTWL
                     IEGYPNAVLRLRFNHFATECSWDHLYVYDGDSIYAPLIAAFSGLIVPERDGNETAPEV
                     TVTSGYALLHFFSDAAYNLTGFNITYNFDMCPNNCSARGECKSSNSSSAVECECSENW
```

KGESCDIPHCTDNCGFPHRGICNASDTRGCSCFPHWQGPGCSIPVPANQSFWTREEYS
DLKLPRASHKAVVNGNIMWVVGGYMFNHSDYSMVLAYDLTSREWLPLNHSVNSVVVRY
GHSLALHKDKIYMYGGKIDSTGNVTNELRVFHIHNESWVLLTPKAKDQYAVVGHSAHI
VTLASGRVVMLVIFGHCPLYGYISVVQEYDLEKNTWSILHTQGALVQGGYGHSSAYDD
RTKALYVHGGYKAFSANKYRLADDLYRYDVDTQMWTILKDSRFFRYLHTAVIVSGTML
VFGGNTHNDTSMSHGAKCFSSDFMAYDIACDRWSVLPRPELHHDVNRFGHSAVLYNST
MYVFGGFNSLLLSDVLVFTSEQCDAHRSEAACVAAGPGIRCLWDTQSSRCTSWELATE
EQAEKLKSECFSKRTLDHDRCDQHTDCYSCTANTNDCHWCNDHCVPVNHSCTEGQISI
AKYESCPKDNPMYYCNKKTSCRSCALDQNCQWEPRNQECIALPENICGNGWHLVGNSC
LKITTAKENYDNAKLSCRNHNAFLASLTSQKKVELVLKQLRLMQSSQSMSKLTLTPWV
GLRKINVSYWCWEDMSPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLNG
SVCERPANHSAKQCRTPCALRTACGECTSSSSECMWCSNMKQCVDSNAYVASFPFGQC
MEWYTMSSCPPENCSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKMPSQA
SAGNVYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSVCEKCEDLTTGKHC
ETCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGDECQLCKVENRYQ
GNPLKGTCYYTLLIDYQFTFSLSQGDDRYYTAINFVATPDEQNRDFDMFINASKKFNL
NITWATSFPAGTQTGEEVPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTWP
IKIQIAFSQHSNFMDLVQFFVTFFSCFLSLLLVAAVVWKIKQSCWASRRREQLLREMQ
QMASRPFASVNVALETDEEPPDLIGGSIKTVPKPIALEPCFGNKAAVLSVFVRLPRGL
GGIPPPGQSGLAVASALVDISQQMPIVYKEKSGAVRNRKQQPPAQPGTCI"

ORIGIN

Alignment Scores:

Pred. No.:	1.82e-187	Length:	4313
Score:	206.00	Matches:	206
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	14.42%	Indels:	0
DB:	10	Gaps:	0

US-09-787-097-12 (1-1429) x AF119821 (1-4313)

```
Qy       1224 ProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPhe 1243
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3693 CCTGTTGTTTCAAAAACCAACATCAAGGAATACAAAGATAGCTTCTCTAATGAGAAATTT 3752

Qy       1244 AspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpPro 1263
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3753 GATTTTCGCAACCATCCAAACATCACTTTCTTTGTTTATGTCAGTAATTTCACTTGGCCC 3812

Qy       1264 IleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePhe 1283
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3813 ATCAAAATTCAGATTGCCTTCTCCCAGCACAGCAACTTTATGGACCTGGTACAGTTCTTC 3872

Qy       1284 ValThrPhePheSerCysPheLeuSerLeuLeuLeuValAlaAlaValValTrpLysIle 1303
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3873 GTGACTTTCTTCAGTTGTTTCCTCTCGCTGCTTCTGGTGGCTGCAGTGGTCTGGAAGATC 3932

Qy       1304 LysGlnSerCysTrpAlaSerArgArgArgGluGlnLeuLeuArgGluMetGlnGlnMet 1323
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3933 AAGCAGAGCTGTTGGGCATCCAGGCGGAGAGAGCAACTTCTTCGGGAGATGCAACAGATG 3992

Qy       1324 AlaSerArgProPheAlaSerValAsnValAlaLeuGluThrAspGluGluProProAsp 1343
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3993 GCCAGCCGCCCCTTTGCTTCTGTAAACGTTGCCTTGGAAACAGATGAAGAACCTCCTGAT 4052

Qy       1344 LeuIleGlyGlySerIleLysThrValProLysProIleAlaLeuGluProCysPheGly 1363
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4053 CTCATTGGGGGAAGTATAAAGACCGTTCCTAAGCCCATTGCCCTGGAGCCCTGCTTTGGT 4112

Qy       1364 AsnLysAlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyGlyIlePro 1383
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4113 AACAAAGCCGCAGTCCTCTCTGTATTCGTGAGGCTCCCTCGAGGACTGGGAGGAATCCCT 4172

Qy       1384 ProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMet 1403
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4173 CCTCCTGGTCAGTCAGGTCTCGCTGTGGCCAGTGCCCTGGTGGACATTTCTCAGCAGATG 4232

Qy       1404 ProIleValTyrLysGluLysSerGlyAlaValArgAsnArgLysGlnGlnProProAla 1423
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4233 CCAATAGTGTACAAGGAGAAGTCAGGAGCTGTAAGAAACCGGAAGCAGCAGCCGCCTGCA 4292

Qy       1424 GlnProGlyThrCysIle 1429
```

```
             ||||||||||||||||||
Db       4293 CAGCCTGGAACCTGCATT 4310
```

RESULT 12
AB038387
LOCUS AB038387 8739 bp mRNA linear ROD 18-JAN-2001
DEFINITION Rattus norvegicus atrn mRNA for attractin, complete cds.
ACCESSION AB038387
VERSION AB038387.1 GI:12275307
KEYWORDS attractin.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (sites)
AUTHORS Kuramoto,T., Kitada,K., Inui,T., Sasaki,Y., Ito,K., Hase,T., Kawaguchi,S., Ogawa,Y., Nakao,K., Barsh,G.S., Nagao,M., Ushijima,T. and Serikawa,T.
TITLE Attractin/mahogany/zitter plays a critical role in myelination of the central nervous system
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (2), 559-564 (2001)
MEDLINE 21143347
PUBMED 11209055
REFERENCE 2 (bases 1 to 8739)
AUTHORS Kuramoto,T., Serikawa,T. and Ushijima,T.
TITLE Direct Submission
JOURNAL Submitted (15-FEB-2000) Takashi Kuramoto, National Cancer Center Research Institute, Carcinogenesis Division; Tsukiji 5-1-1, Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramot@gan2.ncc.go.jp, Tel:81-33542-2511(ex.4521), Fax:81-35565-1753)
FEATURES Location/Qualifiers
source 1. .8739
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="WTC"
/db_xref="taxon:10116"
gene 1. .8739
/gene="atrn"
CDS 61. .4359
/gene="atrn"
/codon_start=1
/product="attractin"
/protein_id="BAB21017.1"
/db_xref="GI:12275308"
/translation="MVAAAAAAEATEARLRGYTTATAAPAGWKERQHRPCAATGAWRP
WPRAGLCLPRVLSRALSPPPLLPLLPLLFSLLLLPLPREAEAAAVAAAVSGSAAAEAK
ECDRPCVNGGRCNPGTGQCVCPTGWVGEQCQHCGGRFRLTGSSGFVTDGPGNYKYKTK
CTWLIEGQPNKIMRLRFNHFATECSWDHLYVYDGDSIYAPLIAAFSGLIVPERDGNET
APEVTVTSGYALLHFFSDAAYNLTGFNITYNFDMCPNNCSGRGECKSSNSSSTVECEC
SENWKGESCDIPHCTDNCGFPHRGICNASDTRGCSCFPHWQGPGCSIPVPANQSFWTR
EEYSDLKLPRASHKAVVNGNIMWVVGGYMFNHSDYSMVLAYDLASREWLSLNHSVNSV
VVRYGHSLALHKDKIYMYGGKIDSTGNVTNELRVFHIHNESWVLLTPKAKDQYAVVGH
SAHIVTLSSGRVVMLVIFGHCPLYGYISVVQEYDLEKNTWSILQTQGALVQGGYGHSS
VYDHRTKALYVHGGYKAFSANKYRLADDLYRYHVDTQMWTILKDSRFFRYLHTAVIVS
GTMLVFGGNTHNDTSMSHGAKCFSSDFMAYDIACDRWSVLPRPELHHDVNRFGHSAVL
HNSTMYVFGGFNSLLLSDVLVFTSEQCDAHRSEAACVAAGPGIRCLWDTQSSRCTSWE
LATEEQAEKLKSECFSKRTLDHDRCDQHTDCYSCTANTNDCHWCNDHCVPVNHSCTEG
QISIAKYDNCPKDNPMYYCNKKTSCRSCALDQNCQWEPRNQECIALPENICGIGWHLV
GNSCLKITTAKENYDNAKLSCRNHNAFLASLTSQKKVEFVLKQLRLMQSSQSTSKLTL
TPWVGLRKINVSYWCWEDMSPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCIN
PLNGSVCERPANHSAKQCRTPCALRTACGECTSSSSECMWCSNMKQCVDSNAYVASFP
FGQCMEWYTMSSCPPENCSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKM
PSHASTGNVYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCEDLTT
GKHCETCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGEECQLCEVE
NRYQGNPLKGTCYYTLLIDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASK
NFNLNITWATSFPAGTQTGEEVPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSN
FTWPIKIQIAFSQHSNFMDLVQFFVTFFSCFLSLLLVAAVVWKIKQSCWASRRREQLL
REMQQMASRPFASVNVALETDEEPPDLIGGSIKTVPKPIALEPCFGNKAAVLSVFVRL
PRGLGGIPPPGQSGLAVASALVDISQQMPIVYKEKSGAVRNRKQQPPAQPGTCI"

ORIGIN

Alignment Scores:

Pred. No.:	3.21e-187	Length:	8739
Score:	206.00	Matches:	568
Percent Similarity:	95.62%	Conservative:	0
Best Local Similarity:	95.62%	Mismatches:	13
Query Match:	14.42%	Indels:	26
DB:	10	Gaps:	0

US-09-787-097-12 (1-1429) x AB038387 (1-8739)

```
Qy        849 SerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2614 TCCAAGCTCACTCTGACTCCATGGGTTGGTCTTCGGAAGATCAATGTATCTTACTGGTGC 2673

Qy        869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2674 TGGGAGGATATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGT 2733

Qy        889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2734 GATGCCGGCTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGACTGAAGGCTGCAACC 2793

Qy        909 CysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSerAlaLysGln 928
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2794 TGCATCAACCCACTCAATGGAAGTGTTTGTGAAAGGCCTGCAAACCACAGTGCCAAGCAG 2853

Qy        929 CysArgThrProCysAlaLeuArgThrAlaCysGlyAsp-CysThrSerGly-SerSerG 948
              |||||||||||||||||||||||||||||||||||||    |||||||    |||||||
Db       2854 TGCCGGACACCCTGTGCCCTGCGGACAGCATGTGGTGA-GTGCACCAGCAG-CAGCTCCG 2911

Qy        948 luCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrValAlaSerP 968
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2912 AATGCATGTGGTGCAGCAACATGAAGCAGTGCGTAGACTCCAATGCCTACGTAGCCTCCT 2971

Qy        968 heProPheGlyGlnCysMetGluTrpTyrThrMetSerThr-CysProProGluAsnCys 987
              ||||||||||||||||||||||||||||||||||||||||    ||||||||||||||||||
Db       2972 TTCCTTTCGGCCAGTGTATGGAGTGGTATACAATGAGCAG-CTGCCCACCTGAAAATTGC 3030

Qy        988 SerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCysThrAsp 1007
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3031 TCTGGCTACTGTACCTGCAGCCATTGCTTGGAGCAGCCAGGCTGTGGTTGGTGTACTGAT 3090

Qy       1008 ProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyProValLysMet 1027
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3091 CCTAGCAATACTGGGAAAGGCAAATGTATTGAGGGAAGCTATAAAGGACCCGTGAAGATG 3150

Qy       1028 ProSerGlnAlaPro--ThrGlyAsnPhe-TyrProGlnProLeuLeuAsnSerSerMet 1046
              ||||||   |||  |||||||||   |||||||||||||||||||||||||||||||||
Db       3151 CCTTCACA-TG-CCTCTACAGGAAATGT-GTACCCACAGCCCCTTCTGAACTCCAGCATG 3207

Qy       1047 CysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnCysAsn 1066
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3208 TGTCTAGAGGACAGCAGATACAACTGGTCTTTTATTCACTGTCCAGCTTGCCAGTGCAAT 3267

Qy       1067 GlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsn-LeuThrThrGl 1086
              ||||||||||||||||||||||||||||||||||||||||||||||||    |||||||||||
Db       3268 GGACACAGCAAGTGCATCAACCAGAGCATCTGTGAGAAGTGTGAGGA-CCTGACCACAGG 3326

Qy       1086 yLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyGlyLysCy 1106
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3327 CAAGCACTGCGAGACCTGCATATCTGGCTTCTATGGTGACCCGACTAATGGAGGCAAATG 3386

Qy       1106 sGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysPh 1126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3387 TCAGCCATGCAAGTGCAACGGGCACGCATCACTGTGCAACACCAACACTGGCAAGTGCTT 3446

Qy       1126 eCysThrThrLysGlyValLysGlyAsp-GluCysGlnLeuCysGluValGluAsnArgT 1146
              ||||||||||||||||||||||||    |||||||||||||||||||||||||||||||||
Db       3447 CTGTACCACCAAGGGCGTCAAGGGAGA-GGAGTGCCAGCTATGTGAGGTAGAAAATCGAT 3505

Qy       1146 yrGlnGlyAsnProLeuArg-GlyThrCysTyrTyrThrLeuLeuIleAspTyrGlnPhe 1165
              ||||||||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db       3506 ACCAGGGAAACCCTCTCAA-AGGAACATGTTACTATACCCTTCTCATTGACTATCAGTTC 3564

Qy       1166 ThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIleAsnPheValAlaThr 1185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3565 ACTTTTAGCCTGTCCCAGGAAGACGACCGCTATTACACAGCCATCAATTTTGTGGCTACT 3624

Qy       1186 ProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeu 1205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3625 CCTGATGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAAAACTTCAACCTC 3684

Qy       1206 AsnIleThrTrpAlaAla-SerPheSer-AlaGlyThrGlnAla-GlyGluGluMet-Pr 1224
              ||||||||||||||||    ||||||    ||||||||||||    |||||||||    ||
Db       3685 AACATCACCTGGGCCAC-CAGTTTCCC-AGCTGGAACCCAGAC-TGGAGAAGAGGT-GCC 3740

Qy       1224 oValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPheAs 1244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3741 TGTTGTTTCAAAAACCAACATCAAGGAATACAAAGATAGCTTCTCTAATGAGAAATTTGA 3800

Qy       1244 pPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpProIl 1264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3801 TTTTCGCAACCATCCAAACATTACTTTCTTTGTTTATGTCAGTAATTTCACTTGGCCCAT 3860

Qy       1264 eLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePheVa 1284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3861 CAAAATTCAAATTGCCTTCTCCCAGCACAGCAATTTTATGGACCTGGTGCAGTTCTTCGT 3920

Qy       1284 lThrPhePheSerCysPheLeuSerLeuLeuLeuValAlaAlaValValTrpLysIleLy 1304
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3921 GACTTTCTTCAGCTGTTTCCTCTCTCTGCTTCTGGTGGCGGCAGTGGTCTGGAAGATCAA 3980

Qy       1304 sGlnSerCysTrpAlaSerArgArgArgGluGlnLeuLeuArgGluMetGlnGlnMetAl 1324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3981 GCAGAGCTGTTGGGCATCCAGGAGGAGAGAGCAACTTCTTCGGGAGATGCAACAGATGGC 4040

Qy       1324 aSerArgProPheAlaSerValAsnValAlaLeuGluThrAspGluGluProProAspLe 1344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4041 CAGCCGCCCTTTTGCTTCTGTAAATGTTGCCTTGGAAACAGATGAAGAACCTCCTGATCT 4100

Qy       1344 uIleGlyGlySerIleLysThrValProLysProIleAlaLeuGluProCysPheGlyAs 1364
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4101 CATTGGGGGAAGTATAAAGACTGTTCCTAAGCCCATTGCCCTGGAGCCCTGCTTTGGTAA 4160

Qy       1364 nLysAlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyGlyIleProPr 1384
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4161 CAAAGCTGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGACTGGGAGGAATCCCTCC 4220

Qy       1384 oProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetPr 1404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4221 TCCTGGTCAGTCAGGTCTTGCTGTGGCCAGTGCCCTGGTGGACATTTCTCAGCAGATGCC 4280

Qy       1404 oIleValTyrLysGluLysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaGl 1424
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4281 AATAGTGTACAAGGAGAAGTCAGGAGCTGTAAGGAACCGGAAGCAGCAGCCCCCTGCACA 4340

Qy       1424 nProGlyThrCysIle 1429
              ||||||||||||||
Db       4341 GCCTGGAACCTGCATT 4356
```

RESULT 13
AR164807
LOCUS AR164807 8827 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 1 from patent US 6274339.
ACCESSION AR164807
VERSION AR164807.1 GI:16238010
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 8827)
AUTHORS Moore,K. and Nagle,D.Lynn.
TITLE Methods and compositions for the diagnosis and treatment of body weight disorders, including obesity
JOURNAL Patent: US 6274339-A 1 14-AUG-2001;
FEATURES Location/Qualifiers
source 1. .8827

```
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:                3.23e-187      Length:          8827
Score:                    206.00         Matches:         576
Percent Similarity:       96.00%         Conservative:    0
Best Local Similarity:    96.00%         Mismatches:      12
Query Match:              14.42%         Indels:          24
DB:                       6              Gaps:            0

US-09-787-097-12 (1-1429) x AR164807 (1-8827)

Qy        842 MetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLys 861
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2790 ATGCAATCATCTCAAAGTATGTCCAAGCTCACTCTGACTCCATGGGTTGGTCTTCGGAAG 2849

Qy        862 IleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGln 881
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2850 ATCAATGTGTCTTACTGGTGCTGGGAGGATATGTCTCCATTCACAAATAGTTTGCTGCAG 2909

Qy        882 TrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThr 901
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2910 TGGATGCCATCTGAGCCCAGTGATGCTGGCTTCTGTGGGATCTTGTCAGAGCCTAGTACT 2969

Qy        902 ArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro 921
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2970 CGGGGATTAAAGGCTGCAACCTGCATCAACCCTCTCAATGGCAGCGTCTGTGAAAGGCCT 3029

Qy        922 AlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAsp 941
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3030 GCAAACCACAGTGCCAAGCAGTGCCGGACACCATGTGCCCTGCGGACAGCGTGTGGCGA- 3088

Qy        942 -CysThrSerGly-SerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspS 961
               |||||||||    ||||||||||||||||||||||||||||||||||||||||||||||
Db       3089 GTGCACTAGCAG-CAGCTCGGAGTGCATGTGGTGCAGTAACATGAAGCAGTGTGTGGACT 3147

Qy        961 erAsnAlaTyrValAlaSerProProPheGlyGlnCysMetGluTrpTyrThrMetSerT 981
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3148 CCAATGCCTACGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCA 3207

Qy        981 hr-CysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPro 1000
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3208 G-CTGCCCACCTGAAAATTGCTCTGGCTACTGTACCTGCAGCCATTGCTTGGAGCAGCCA 3266

Qy       1001 GlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySer 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3267 GGCTGTGGTTGGTGTACTGATCCTAGCAATACTGGGAAAGGAAAATGTATTGAGGGCAGC 3326

Qy       1021 TyrLysGlyProValLysMetProSerGlnAlaProThr--GlyAsnPhe-TyrProGln 1039
              |||||||||||||||||||||||||||||||||||||||  ||||||  |||||||||||||
Db       3327 TATAAAGGACCTGTGAAGATGCCGTCACAGGCCTC-TG-CAGGAAATGT-GTATCCACAG 3383

Qy       1040 ProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHis 1059
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3384 CCCCTTCTGAACTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCAC 3443

Qy       1060 CysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLys 1079
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3444 TGTCCAGCTTGCCAGTGCAACGGACACAGCAAATGCATCAACCAGAGTATCTGTGAGAAG 3503

Qy       1080 CysGluAsn-LeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAs 1099
              ||||||    ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3504 TGTGAGGA-CCTGACCACGGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTATGGTGA 3562

Qy       1099 pProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAs 1119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3563 CCCGACTAATGGAGGCAAATGTCAGCCATGCAAGTGCAATGGGCACGCATCACTGTGCAA 3622

Qy       1119 nThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLe 1139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3623 CACCAACACCGGCAAGTGCTTCTGTACCACCAAAGGTGTCAAGGGGGACGAGTGCCAGCT 3682

Qy       1139 uCysGluValGluAsnArgTyrGlnGlyAsnProLeuArg-GlyThrCysTyrTyrThrL 1159
              ||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||
Db       3683 ATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAA-AGGAACATGCTACTATACCC 3741

Qy       1159 euLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrA 1179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3742 TTCTCATTGACTATCAGTTCACCTTTAGCCTGTCCCAGGAAGACGACCGCTACTACACAG 3801

Qy       1179 laIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnA 1199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3802 CCATCAACTTTGTGGCTACTCCTGATGAACAAAACAGGGATTTGGACATGTTCATCAATG 3861

Qy       1199 laSerLysAsnPheAsnLeuAsnIleThrTrpAlaAla-SerPheSer-AlaGlyThrGl 1218
              ||||||||||||||||||||||||||||||||||||  ||||||    ||||||||||||
Db       3862 CCTCCAAAAACTTCAACCTCAACATCACCTGGGCCAC-CAGCTTCCC-AGCCGGAACCCA 3919

Qy       1218 nAla-GlyGluGluMet-ProValValSerLysThrAsnIleLysGluTyrLysAspSer 1237
              |   |||||||||    ||||||||||||||||||||||||||||||||||||||||||||
Db       3920 GAC-TGGAGAAGAGGT-GCCTGTTGTTTCAAAAACCAACATCAAGGAATACAAAGATAGC 3977

Qy       1238 PheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrVal 1257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3978 TTCTCTAATGAGAAATTTGATTTTCGCAACCATCCAAACATCACTTTCTTTGTTTATGTC 4037

Qy       1258 SerAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMet 1277
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4038 AGTAATTTCACTTGGCCCATCAAAATTCAGATTGCCTTCTCCCAGCACAGCAACTTCATG 4097

Qy       1278 AspLeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuLeuValAla 1297
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4098 GACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCGCTGCTTCTGGTGGCT 4157

Qy       1298 AlaValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeuLeu 1317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4158 GCAGTGGTCTGGAAGATCAAGCAGAGCTGTTGGGCATCCAGGCGGAGAGAGCAACTTCTT 4217

Qy       1318 ArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThr 1337
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4218 CGGGAGATGCAACAGATGGCCAGCCGCCCCTTTGCTTCTGTAAACGTTGCCTTGGAAACA 4277

Qy       1338 AspGluGluProProAspLeuIleGlyGlySerIleLysThrValProLysProIleAla 1357
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4278 GATGAGGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCA 4337

Qy       1358 LeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProArg 1377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4338 CTGGAGCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGA 4397

Qy       1378 GlyLeuGlyGlyIleProProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuVal 1397
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4398 GGCCTGGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTG 4457

Qy       1398 AspIleSerGlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAsnArg 1417
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4458 GACATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGG 4517

Qy       1418 LysGlnGlnProProAlaGlnProGlyThrCysIle 1429
              |||||||||||||||||||||||||||||||||||||
Db       4518 AAGCAGCAGCCCCCTGCACAGCCTGGGACCTGCATC 4553

RESULT 14
AB093245
LOCUS       AB093245               5683 bp    mRNA    linear   ROD 07-FEB-2003
DEFINITION  Mus musculus mRNA for mKIAA0548 protein.
ACCESSION   AB093245
VERSION     AB093245.1  GI:26006172
KEYWORDS    FLI_CDNA.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
```

```
  AUTHORS   Okazaki,N., Kikuno,R., Ohara,R., Inamoto,S., Hara,Y., Nagase,T.,
            Ohara,O. and Koga,H.
  TITLE     Prediction of the coding sequences of mouse homologues of KIAA
            gene: I. The complete nucleotide sequences of 100 mouse
            KIAA-homologous cDNAs identified by screening of terminal sequences
            of cDNA clones randomly sampled from size-fractionated libraries
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 5683)
  AUTHORS   Okazaki,N., Kikuno,R., Nagase,T., Ohara,O. and Koga,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-OCT-2002) Hisashi Koga, Kazusa DNA Research
            Institute, Laboratory for Genome Informatics; 2-6-7
            Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan
            (E-mail:mouse@kazusa.or.jp, Tel:81-438-52-3919, Fax:81-438-52-3918)
COMMENT     The CREATE program supported by Japan science and technology
            corporation; cDNA full insert sequencing: Kazusa DNA Research
            Institute; cDNA library construction, clone selection and 5'- &
            3'-end one pass sequencing.
FEATURES             Location/Qualifiers
     source          1. .5683
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="mbg03041"
                     /tissue_type="brain"
                     /dev_stage="adult"
                     /note="vector:modified pBC SK+"
     gene            1. .5683
                     /gene="mKIAA0548"
     CDS             <1. .3181
                     /gene="mKIAA0548"
                     /note="CDS is predicted by in silico analysis. Start codon
                     is not identified."
                     /codon_start=2
                     /evidence=not_experimental
                     /product="mKIAA0548 protein"
                     /protein_id="BAC41429.1"
                     /db_xref="GI:26006173"
                     /translation="AYDLTSREWLPLNHSVNSVVVRYGHSLALHKDKIYMYGGKIDST
                     GNVTNELRVFHIHNESWVLLTPKAKDQYAVVGHSAHIVTLASGRVVMLVIFGHCPLYG
                     YISVVQEYDLEKNTWSILHTQGALVQGGYGHSSVYDDRTKALYVHGGYKAFSANKYRL
                     ADDLYRYDVDTQMWTILKDSRFFRYLHTAVIVSGTMLVFGGNTHNDTSMSHGAKCFSS
                     DFMAYDIACDRWSVLPRPELHHDVNRFGHSAVLYNSTMYVFGGFNSLLLSDVLVFTLE
                     QCDAHRSEAACVAAGPGIRCLWDTQSSRCTSWELATEEQAEKLKSECFSKRTLDHDRC
                     DQHTDCYSCTANTNDCHWCNDHCVPVNHSCTEGQISIAKYESCPKDNPMYYCNKKTSC
                     RSCALDQNCQWEPRNQECIALPENICGNGWHLVGNSCLKITTAKENYDNAKLSCRNHN
                     AFLASLTSQKKVEFVLKQLRLMQSSQSMSKLTLTPWVGLRKINVSYWCWEDMSPFTNS
                     LLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLNGSVCERPANHSAKQCRTPCALR
                     TACGECTSSSSECMWCSNMKQCVDSNAYVASFPFGQCMEWYTMSSCPPENCSGYCTCS
                     HCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKMPSQASAGNVYPQPLLNSSMCLEDSR
                     YNWSFIHCPACQCNGHSKCINQSVCEKCEDLTTGKHCETCISGFYGDPTNGGKCQPCK
                     CNGHASLCNTNTGKCFCTTKGVKGDECQLCEVENRYQGNPLKGTCYYTLLIDYQFTFS
                     LSQEDDRYYTAINFVATPDEQNRDLDMFINASKNFNLNITWATSFPAGTQTGEEVPVV
                     SKTTIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTWPIKIQIAFSQHSNFMDLVQFFV
                     TFFSCFLSLLLVAAVVWKIKQSCWASRRREQLLREMQQMASRPFASVNVALETDEEPP
                     DLIGGSIKTVPKPIALEPCFGNKAAVLSVFVRLPRGLGGIPPPGQSGLAVASALVDIS
                     QQMPIVYKEKSGAVRNRKQQPPAQPGTCI"
ORIGIN

Alignment Scores:
Pred. No.:                1.25e-180       Length:          5683
Score:                    199.00          Matches:         574
Percent Similarity:       95.35%          Conservative:    0
Best Local Similarity:    95.35%          Mismatches:      14
Query Match:              13.93%          Indels:          28
DB:                       10              Gaps:            0

US-09-787-097-12 (1-1429) x AB093245 (1-5683)

Qy        842 MetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLys 861
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1415 ATGCAATCATCTCAAAGTATGTCCAAGCTCACTCTGACTCCATGGGTTGGTCTTCGGAAG 1474

Qy        862 IleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGln 881
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1475 ATCAATGTGTCTTACTGGTGCTGGGAGGATATGTCTCCATTCACAAATAGTTTGCTGCAG 1534

Qy        882 TrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThr 901
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1535 TGGATGCCATCTGAGCCCAGTGATGCTGGCTTCTGTGGGATCTTGTCAGAGCCTAGTACT 1594

Qy        902 ArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro 921
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1595 CGGGGATTAAAGGCTGCAACCTGCATCAACCCTCTCAATGGCAGCGTCTGTGAAAGGCCT 1654

Qy        922 AlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAsp 941
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1655 GCAAACCACAGTGCCAAGCAGTGCCGGACACCATGTGCCCTGCGGACAGCGTGTGGCGA- 1713

Qy        942 -CysThrSerGly-SerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspS 961
               ||||||||    ||||||||||||||||||||||||||||||||||||||||||||||
Db       1714 GTGCACTAGCAG-CAGCTCGGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACT 1772

Qy        961 erAsnAlaTyrValAlaSerProProPheGlyGlnCysMetGluTrpTyrThrMetSerT 981
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1773 CCAATGCCTACGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCA 1832

Qy        981 hr-CysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPro 1000
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1833 G-CTGCCCACCTGAAAATTGCTCTGGCTATTGTACCTGCAGCCATTGCTTGGAGCAGCCA 1891

Qy       1001 GlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySer 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1892 GGCTGTGGTTGGTGTACTGATCCTAGCAATACTGGGAAAGGAAAATGTATTGAGGGCAGC 1951

Qy       1021 TyrLysGlyProValLysMetProSerGlnAlaProThr--GlyAsnPhe-TyrProGln 1039
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1952 TATAAAGGACCTGTGAAGATGCCGTCACAGGCCTC-TG-CAGGAAATGT-GTATCCACAG 2008

Qy       1040 ProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHis 1059
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2009 CCCCTTCTGAACTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCAC 2068

Qy       1060 CysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIle-CysGluLy 1079
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2069 TGTCCAGCTTGCCAGTGCAACGGACACAGCAAATGCATCAACCAGAGTGT-CTGTGAGAA 2127

Qy       1079 sCysGluAsn-LeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyA 1099
              |||||||    ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2128 GTGTGAGGA-CCTGACCACGGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTATGGTG 2186

Qy       1099 spProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysA 1119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2187 ACCCGACTAATGGAGGCAAATGTCAGCCATGCAAGTGCAATGGGCACGCATCACTGTGCA 2246

Qy       1119 snThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnL 1139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2247 ACACCAACACCGGCAAGTGCTTCTGTACCACCAAAGGTGTCAAGGGGGACGAGTGCCAGC 2306

Qy       1139 euCysGluValGluAsnArgTyrGlnGlyAsnProLeuArg-GlyThrCysTyrTyrThr 1158
              ||||||||||||||||||||||||||||||||||||||||    ||||||||||||||||||
Db       2307 TATGTGAGGTAGAAAATCGATACCAGGGAAATCCTCTCAA-AGGAACATGCTACTATACC 2365

Qy       1159 LeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThr 1178
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2366 CTTCTCATTGACTATCAGTTCACCTTTAGCCTGTCCCAGGAAGACGACCGCTACTACACA 2425

Qy       1179 AlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsn 1198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2426 GCCATCAACTTTGTGGCTACTCCTGATGAACAAAACAGGGATTTGGACATGTTCATCAAT 2485

Qy       1199 AlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAla-SerPheSer-AlaGlyThrG 1218
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2486 GCCTCCAAAAACTTCAACCTCAACATCACCTGGGCCAC-CAGCTTCCC-AGCCGGAACCC 2543

Qy       1218 lnAla-GlyGluGluMet-ProValValSerLysThrAsn-IleLysGluTyrLysAspS 1237
```

```
Db       2544 AGAC-TGGAGAAGAGGT-GCCTGTTGTTTCAAAAACCAC-CATCAAGGAATACAAAGATA 2600

Qy       1237 erPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrV 1257

Db       2601 GCTTCTCTAATGAGAAATTTGATTTTCGCAACCATCCAAACATCACTTTCTTTGTTTATG 2660

Qy       1257 alSerAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheM 1277

Db       2661 TCAGTAATTTCACTTGGCCCATCAAAATTCAGATTGCCTTCTCCCAGCACAGCAACTTCA 2720

Qy       1277 etAspLeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuLeuValA 1297

Db       2721 TGGACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCGCTGCTTCTGGTGG 2780

Qy       1297 laAlaValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeuL 1317

Db       2781 CTGCAGTGGTCTGGAAGATCAAGCAGAGCTGTTGGGCATCCAGGCGGAGAGAGCAACTTC 2840

Qy       1317 euArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluT 1337

Db       2841 TTCGGGAGATGCAACAGATGGCCAGCCGCCCCTTTGCTTCTGTAAACGTTGCCTTGGAAA 2900

Qy       1337 hrAspGluGluProProAspLeuIleGlyGlySerIleLysThrValProLysProIleA 1357

Db       2901 CAGATGAAGAACCTCCTGATCTCATTGGGGGAAGTATAAAGACCGTTCCTAAGCCCATTG 2960

Qy       1357 laLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProA 1377

Db       2961 CCCTGGAGCCCTGCTTTGGTAACAAAGCCGCAGTCCTCTCTGTATTCGTGAGGCTCCCTC 3020

Qy       1377 rgGlyLeuGlyGlyIleProProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuV 1397

Db       3021 GAGGACTGGGAGGAATCCCTCCTCCTGGTCAGTCAGGTCTCGCTGTGGCCAGTGCCCTGG 3080

Qy       1397 alAspIleSerGlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAsnA 1417

Db       3081 TGGACATTTCTCAGCAGATGCCAATAGTGTACAAGGAGAAGTCAGGAGCTGTAAGAAACC 3140

Qy       1417 rgLysGlnGlnProProAlaGlnProGlyThrCysIle 1429

Db       3141 GGAAGCAGCAGCCGCCTGCACAGCCTGGAACCTGCATT 3178
```

```
RESULT 15
AF116897
LOCUS       AF116897                6083 bp    mRNA    linear   ROD 20-MAR-1999
DEFINITION  Mus musculus mahogany protein mRNA, complete cds.
ACCESSION   AF116897
VERSION     AF116897.1  GI:4454560
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 6083)
  AUTHORS   Nagle,D.L., McGrail,S.H., Vitale,J., Woolf,E.A., Dussault,B.J. Jr.,
            DiRocco,L., Holmgren,L., Montagno,J., Bork,P., Huszar,D.,
            Fairchild-Huntress,V., Ge,P., Keilty,J., Ebeling,C., Baldini,L.,
            Gilchrist,J., Burn,P., Carlson,G.A. and Moore,K.J.
  TITLE     The mahogany protein is a receptor involved in suppression of
            obesity
  JOURNAL   Nature 398 (6723), 148-152 (1999)
  MEDLINE   99184159
   PUBMED   10086355
REFERENCE   2  (bases 1 to 6083)
  AUTHORS   Nagle,D.L., McGrail,S.H., Vitale,J., Woolf,E.A., Dussault,B.J. Jr.,
            DiRocco,L., Holmgren,L., Montagno,J., Bork,P., Huszar,D.,
            Fairchild-Huntress,V., Ge,P., Keilty,J., Ebeling,C., Baldini,L.,
            Gilchrist,J., Burn,P., Carlson,G.A. and Moore,K.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-DEC-1998) Sequence Analysis, Millennium
            Pharmaceuticals, 640 Memorial Drive, Cambridge, MA 02139, USA
FEATURES             Location/Qualifiers
     source          1. .6083
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
     CDS             70. .4488
                     /note="MG"
                     /codon_start=1
                     /product="mahogany protein"
                     /protein_id="AAD20947.1"
                     /db_xref="GI:4454561"
                     /translation="MVAVAAAAATEARLRGSTTTTAAPAGRKGRQHRPCTATGAWRPG
                     PRARLCLPRVLSRALPPPPLLPLLFSLLLLPLPREAEAAAVAAAVSGSAAAEAKECDR
                     PCVNGGRCNPGTGQCVCPTGWVGEQCQHCGGRFRLTGSSGFVTDGPGNYKYKTKCTWL
                     IEGQPNRIMRLRFNHFATECSWDHLYVYDGDSIYAPLIAAFSGLIVPERDGNETAPEV
                     TVTSGYALLHFFSDAAYNLTGFNITYNFDMCPNNCSGRGECKSSNSSSAVECECSENW
                     KGESCDIPHCTDNCGFPHRGICNASDTRGCSCFPHWQGPGCSIPVPANQSFWTREEYS
                     DLKLPRASHKAVVNGNIMWVVGGYMFNHSDYSMVLAYDLTSREWLPLNHSVNSVVVRY
                     GHSLALHKDKIYMYGGKIDSTGNVTNELRVFHIHNESWVLLTPKAKDQYAVVGHSAHI
                     VTLASGRVVMLVIFGHCPLYGYISVVQEYDLEKNTWSILHTQGALVQGGYGHSSVYDD
                     RTKALYVHGGYKAFSANKYRLADDLYRYDVDTQMWTILKDSRFFRYLHTAVIVSGTML
                     VFGGNTHNDTSMSHGAKCFSSDFMAYDIACDRWSVLPRPELHHDVNRFGHSAVLYNST
                     MYVFGGFNSLLLSDVLVFTSEQCDAHRSEAACVAAGPGIRCLWDTQSSRCTSWELATE
                     EQAEKLKSECFSKRTLDHDRCDQHTDCYSCTANTNDCHWCNDHCVPVNHSCTEGQISI
                     AKYESCPKDNPMYYCNKKTSCRSCALDQNCQWEPRNQECIALPENICGNGWHLVGNSC
                     LKITTAKENYDNAKLSCRNHNAFLASLTSQKKVEFVLKQLRLMQSSQSMSKLTLTPWV
                     GLRKINVSYWCWEDMSPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLNG
                     SVCERPANHSAKQCRTPCALRTACGECTSSSSECMWCSNMKQCVDSNAYVASFPFGQC
                     MEWYTMSSCPPENCSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKMPSQA
                     SAGNVYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCEDLTTGKHC
                     ETCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGDECQLCEVENRYQ
                     GNPLKGTCYYTLLIDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASKNFNL
                     NITWATSFPAGTQTGEEVPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTWP
                     IKIQIAFSQHSNFMDLVQFFVTFFSCFLSLLLVAAVVWKIKQSCWASRRREQLLREMQ
                     QMASRPFASVNVALETDEEPPDLIGGSIKTVPKPIALEPCFGNKAAVLSVFVRLPRGL
                     GGIPPPGQSRSRCGQCPGGHFSADANSVQGEVRSCKKPEAAAACTAWNLHLILGQEFA
                     LTEGVWSTKLSTGEAAGRKQAWGRRLGTLEASNSCA"
ORIGIN

Alignment Scores:
Pred. No.:                7.86e-175       Length:          6083
Score:                    193.00          Matches:         576
Percent Similarity:       95.84%          Conservative:    0
Best Local Similarity:    95.84%          Mismatches:      12
Query Match:              13.51%          Indels:          25
DB:                       10              Gaps:            0

US-09-787-097-12 (1-1429) x AF116897 (1-6083)

Qy        842 MetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLys 861

Db       2590 ATGCAATCATCTCAAAGTATGTCCAAGCTCACTCTGACTCCATGGGTTGGTCTTCGGAAG 2649

Qy        862 IleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGln 881

Db       2650 ATCAATGTGTCTTACTGGTGCTGGGAGGATATGTCTCCATTCACAAATAGTTTGCTGCAG 2709

Qy        882 TrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThr 901

Db       2710 TGGATGCCATCTGAGCCCAGTGATGCTGGCTTCTGTGGGATCTTGTCAGAGCCTAGTACT 2769

Qy        902 ArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro 921

Db       2770 CGGGGATTAAAGGCTGCAACCTGCATCAACCCTCTCAATGGCAGCGTCTGTGAAAGGCCT 2829

Qy        922 AlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAsp 941

Db       2830 GCAAACCACAGTGCCAAGCAGTGCCGGACACCATGTGCCCTGCGGACAGCGTGTGGCGA- 2888

Qy        942 -CysThrSerGly-SerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspS 961

Db       2889 GTGCACTAGCAG-CAGCTCGGAGTGCATGTGGTGCAGTAACATGAAGCAGTGTGTGGACT 2947

Qy        961 erAsnAlaTyrValAlaSerProProGlyGlnCysMetGluTrpTyrThrMetSerT 981
```

```
Db      2948 CCAATGCCTACGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCA 3007

Qy       981 hr-CysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPro 1000

Db      3008 G-CTGCCCACCTGAAAATTGCTCTGGCTACTGTACCTGCAGCCATTGCTTGGAGCAGCCA 3066

Qy      1001 GlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySer 1020

Db      3067 GGCTGTGGTTGGTGTACTGATCCTAGCAATACTGGGAAAGGAAAATGTATTGAGGGCAGC 3126

Qy      1021 TyrLysGlyProValLysMetProSerGlnAlaProThr--GlyAsnPhe-TyrProGln 1039

Db      3127 TATAAAGGACCTGTGAAGATGCCGTCACAGGCCTC-TG-CAGGAAATGT-GTATCCACAG 3183

Qy      1040 ProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHis 1059

Db      3184 CCCCTTCTGAACTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCAC 3243

Qy      1060 CysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLys 1079

Db      3244 TGTCCAGCTTGCCAGTGCAACGGACACAGCAAATGCATCAACCAGAGTATCTGTGAGAAG 3303

Qy      1080 CysGluAsn-LeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAs 1099

Db      3304 TGTGAGGA-CCTGACCACGGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTATGGTGA 3362

Qy      1099 pProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAs 1119

Db      3363 CCCGACTAATGGAGGCAAATGTCAGCCATGCAAGTGCAATGGGCACGCATCACTGTGCAA 3422

Qy      1119 nThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLe 1139

Db      3423 CACCAACACCGGCAAGTGCTTCTGTACCACCAAAGGTGTCAAGGGGGACGAGTGCCAGCT 3482

Qy      1139 uCysGluValGluAsnArgTyrGlnGlyAsnProLeuArg-GlyThrCysTyrTyrThrL 1159

Db      3483 ATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAA-AGGAACATGCTACTATACCC 3541

Qy      1159 euLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrA 1179

Db      3542 TTCTCATTGACTATCAGTTCACCTTTAGCCTGTCCCAGGAAGACGACCGCTACTACACAG 3601

Qy      1179 laIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnA 1199

Db      3602 CCATCAACTTTGTGGCTACTCCTGATGAACAAAACAGGGATTTGGACATGTTCATCAATG 3661

Qy      1199 laSerLysAsnPheAsnLeuAsnIleThrTrpAlaAla-SerPheSer-AlaGlyThrGl 1218

Db      3662 CCTCCAAAAACTTCAACCTCAACATCACCTGGGCCAC-CAGCTTCCC-AGCCGGAACCCA 3719

Qy      1218 nAla-GlyGluGluMet-ProValValSerLysThrAsnIleLysGluTyrLysAspSer 1237

Db      3720 GAC-TGGAGAAGAGGT-GCCTGTTGTTTCAAAAACCAACATCAAGGAATACAAAGATAGC 3777

Qy      1238 PheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrVal 1257

Db      3778 TTCTCTAATGAGAAATTTGATTTTCGCAACCATCCAAACATCACTTTCTTTGTTTATGTC 3837

Qy      1258 SerAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMet 1277

Db      3838 AGTAATTTCACTTGGCCCATCAAAATTCAGATTGCCTTCTCCCAGCACAGCAACTTCATG 3897

Qy      1278 AspLeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuLeuValAla 1297

Db      3898 GACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCGCTGCTTCTGGTGGCT 3957

Qy      1298 AlaValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeuLeu 1317

Db      3958 GCAGTGGTCTGGAAGATCAAGCAGAGCTGTTGGGCATCCAGGCGGAGAGAGCAACTTCTT 4017

Qy      1318 ArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThr 1337

Db      4018 CGGGAGATGCAACAGATGGCCAGCCGCCCCTTTGCTTCTGTAAACGTTGCCTTGGAAACA 4077

Qy      1338 AspGluGluProProAspLeuIleGlyGlySerIleLysThrValProLysProIleAla 1357

Db      4078 GATGAAGAACCTCCTGATCTCATTGGGGGAAGTATAAAGACCGTTCCTAAGCCCATTGCC 4137

Qy      1358 LeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProArg 1377

Db      4138 CTGGAGCCCTGCTTTGGTAACAAAGCCGCAGTCCTCTCTGTATTCGTGAGGCTCCCTCGA 4197

Qy      1378 GlyLeuGlyGlyIleProProProGlyGlnSer-GlyLeuAlaValAlaSerAlaLeuVa 1397

Db      4198 GGACTGGGAGGAATCCCTCCTCCTGGTCAGTCAAGGTCTCGCTGTGGCCAGTGCCCTGGT 4257

Qy      1397 lAspIleSerGlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAsnAr 1417

Db      4258 GGACATTTCTCAGCAGATGCCAATAGTGTACAAGGAGAAGTCAGGAGCTGTAAGAAACCG 4317

Qy      1417 gLysGlnGlnProProAlaGlnProGlyThrCysIle 1429

Db      4318 GAAGCAGCAGCCGCCTGCACAGCCTGGAACCTGCATT 4354
```

RESULT 16
AX778545

LOCUS	AX778545	543 bp DNA linear PAT 14-JUL-2003
DEFINITION	Sequence 702 from Patent WO03039443.	
ACCESSION	AX778545	
VERSION	AX778545.1 GI:32695539	
KEYWORDS	.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1	
AUTHORS	Haferlach,T., Schoch,C., Kern,W., Kohlmann,A., Schnittger,S., Dugas,M., Eils,R., Brors,B. and Mergenthaler,S.	
TITLE	Novel genetic markers for leukemias	
JOURNAL	Patent: WO 03039443-A 702 15-MAY-2003; Deutsches Krebsforschungszentrum (DE) ; Ludwig-Maximilian-Universitaet Muenchen (DE) ; Haferlach, Torsten, PD Dr. Dr. (DE) ; Schoch, Claudia (DE) ; Kern, Wolfgang (DE)	

FEATURES Location/Qualifiers
source 1. .543
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.:	4.05e-164	Length:	543
Score:	181.00	Matches:	181
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.67%	Indels:	0
DB:	6	Gaps:	0

US-09-787-097-12 (1-1429) x AX778545 (1-543)

```
Qy      1084 ThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGly 1103

Db         1 ACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGA 60

Qy      1104 GlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGly 1123

Db        61 GGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGC 120

Qy      1124 LysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGlu 1143

Db       121 AAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAA 180

Qy      1144 AsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyr 1163

Db       181 AATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTAT 240
```

```
Qy       1164 GlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIleAsnPheVal 1183
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTG 300

Qy       1184 AlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPhe 1203
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTC 360

Qy       1204 AsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMet 1223
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATG 420

Qy       1224 ProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPhe 1243
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTT 480

Qy       1244 AspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpPro 1263
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCC 540

Qy       1264 Ile 1264
             |||
Db        541 ATC 543
```

RESULT 17
BX004762_0/c
WPCOMMENT
Sequence split into 5 fragments LOCUS BX004762 Accession BX004762

Fragment Name	Begin	End
BX004762_0	1	110000
BX004762_1	100001	210000
BX004762_2	200001	310000
BX004762_3	300001	410000
BX004762_4	400001	442165

LOCUS BX004762 442165 bp DNA linear HTG 12-APR-2003
DEFINITION Mus musculus chromosome X clone RP23-146C4, WORKING DRAFT SEQUENCE, 4 unordered pieces.
ACCESSION BX004762
VERSION BX004762.6 GI:29824756
KEYWORDS HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 442165)
AUTHORS Chapman,J.
TITLE Direct Submission
JOURNAL Submitted (11-APR-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT On Apr 13, 2003 this sequence version replaced gi:29125168.
Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.
---------- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
---------- Project Information
Center project name: bM146C4
---------- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 229053 bases at least Q40
Consensus quality: 229393 bases at least Q30
Consensus quality: 229535 bases at least Q20
Insert size: 441865; sum-of-contigs
Insert size: 188914; 6.6% error; agarose-fp
Quality coverage: 3.34x in Q20 bases; sum-of-contigs Quality coverage: 7.96x in Q20 bases; agarose-fp
----------.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 48919: contig of 48919 bp in length
* 48920 49019: gap of 100 bp
* 49020 130781: contig of 81762 bp in length
* 130782 130881: gap of 100 bp
* 130882 229964: contig of 99083 bp in length
* 229965 230064: gap of 100 bp
* 230065 442165: contig of 212101 bp in length.

FEATURES	Location/Qualifiers
source	1. .442165
	/organism="Mus musculus"
	/mol_type="genomic DNA"
	/db_xref="taxon:10090"
	/chromosome="X"
	/clone="RP23-146C4"
	/clone_lib="RPCI-23"
misc_feature	1. .48919
	/note="assembly_fragment:00147
	fragment_chain:1"
misc_feature	49020. .130781
	/note="assembly_fragment:00500
	fragment_chain:1"
misc_feature	130882. .229964
	/note="assembly_fragment:00870
	fragment_chain:1"
misc_feature	230065. .442165
	/note="assembly_fragment:02147"

ORIGIN

Alignment Scores:

Pred. No.:	3.59e-113	Length:	110000
Score:	130.00	Matches:	130
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	9.10%	Indels:	0
DB:	2	Gaps:	0

US-09-787-097-12 (1-1429) x BX004762_0 (1-110000)

```
Qy       1259 AsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAsp 1278
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     101355 AATTTCACTTGGCCCATCAAAATTCAGATTGCCTTCTCCCAGCACAGCAACTTCATGGAC 101296

Qy       1279 LeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuLeuValAlaAla 1298
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     101295 CTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCGCTGCTTCTGGTGGCTGCA 101236

Qy       1299 ValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeuLeuArg 1318
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     101235 GTGGTCTGGAAGATCAAGCAGAGCTGTTGGGCATCCAGGCGGAGAGAGCAACTTCTTCGG 101176

Qy       1319 GluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThrAsp 1338
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     101175 GAGATGCAACAGATGGCCAGCCGCCCCTTTGCTTCTGTAAACGTTGCCTTGGAAACAGAT 101116

Qy       1339 GluGluProProAspLeuIleGlyGlySerIleLysThrValProLysProIleAlaLeu 1358
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     101115 GAAGAACCTCCTGATCTCATTGGGGGAAGTATAAAGACCGTTCCTAAGCCCATTGCCCTG 101056

Qy       1359 GluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProArgGly 1378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     101055 GAGCCCTGCTTTGGTAACAAAGCCGCAGTCCTCTCTGTATTCGTGAGGCTCCCTCGAGGA 100996

Qy       1379 LeuGlyGlyIleProProProGlyGlnSer 1388
             |||||||||||||||||||||||||||||
```

```
Db     100995 CTGGGAGGAATCCCTCCTCCTGGTCAGTCA 100966

RESULT 18
BX004762_1/c
WPCOMMENT
  Sequence split into 5 fragments  LOCUS BX004762 Accession BX004762
     Fragment Name      Begin     End
     BX004762_0             1     110000
     BX004762_1        100001     210000
     BX004762_2        200001     310000
     BX004762_3        300001     410000
     BX004762_4        400001     442165
Continuation (2 of 5) of BX004762 from base 100001 (BX004762 Mus musculus chromosome X c

Alignment Scores:
Pred. No.:                3.59e-113      Length:         110000
Score:                    130.00         Matches:        130
Percent Similarity:       100.00%        Conservative:   0
Best Local Similarity:    100.00%        Mismatches:     0
Query Match:              9.10%          Indels:         0
DB:                       2              Gaps:           0

US-09-787-097-12 (1-1429) x BX004762_1 (1-110000)

Qy       1259 AsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAsp 1278
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1355 AATTTCACTTGGCCCATCAAAATTCAGATTGCCTTCTCCCAGCACAGCAACTTCATGGAC 1296

Qy       1279 LeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuLeuValAlaAla 1298
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1295 CTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCGCTGCTTCTGGTGGCTGCA 1236

Qy       1299 ValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeuLeuArg 1318
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1235 GTGGTCTGGAAGATCAAGCAGAGCTGTTGGGCATCCAGGCGGAGAGAGCAACTTCTTCGG 1176

Qy       1319 GluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThrAsp 1338
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1175 GAGATGCAACAGATGGCCAGCCGCCCCTTTGCTTCTGTAAACGTTGCCTTGGAAACAGAT 1116

Qy       1339 GluGluProProAspLeuIleGlyGlySerIleLysThrValProLysProIleAlaLeu 1358
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1115 GAAGAACCTCCTGATCTCATTGGGGGAAGTATAAAGACCGTTCCTAAGCCCATTGCCCTG 1056

Qy       1359 GluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProArgGly 1378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1055 GAGCCCTGCTTTGGTAACAAAGCCGCAGTCCTCTCTGTATTCGTGAGGCTCCCTCGAGGA 996

Qy       1379 LeuGlyGlyIleProProProGlyGlnSer 1388
             |||||||||||||||||||||||||||||
Db        995 CTGGGAGGAATCCCTCCTCCTGGTCAGTCA 966

RESULT 19
AL807401
LOCUS       AL807401               110261 bp    DNA     linear   ROD 31-OCT-2002
DEFINITION  Mouse DNA sequence from clone RP23-133C12 on chromosome X, complete
            sequence.
ACCESSION   AL807401
VERSION     AL807401.11  GI:24474469
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 110261)
  AUTHORS   Howden,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-OCT-2002) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Nov 1, 2002 this sequence version replaced gi:24395101.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-133C12 is
            from the RPCI-23 Mouse PAC Library
            constructed by the group of Pieter de Jong.
            For further details see http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6.
FEATURES             Location/Qualifiers
     source          1. .110261
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="X"
                     /clone="RP23-133C12"
                     /clone_lib="RPCI-23"
ORIGIN

Alignment Scores:
Pred. No.:                3.59e-113      Length:         110261
Score:                    130.00         Matches:        130
Percent Similarity:       100.00%        Conservative:   0
Best Local Similarity:    100.00%        Mismatches:     0
Query Match:              9.10%          Indels:         0
DB:                       10             Gaps:           0

US-09-787-097-12 (1-1429) x AL807401 (1-110261)

Qy       1259 AsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAsp 1278
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      33174 AATTTCACTTGGCCCATCAAAATTCAGATTGCCTTCTCCCAGCACAGCAACTTCATGGAC 33233

Qy       1279 LeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuLeuValAlaAla 1298
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      33234 CTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCGCTGCTTCTGGTGGCTGCA 33293

Qy       1299 ValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeuLeuArg 1318
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      33294 GTGGTCTGGAAGATCAAGCAGAGCTGTTGGGCATCCAGGCGGAGAGAGCAACTTCTTCGG 33353

Qy       1319 GluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThrAsp 1338
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      33354 GAGATGCAACAGATGGCCAGCCGCCCCTTTGCTTCTGTAAACGTTGCCTTGGAAACAGAT 33413

Qy       1339 GluGluProProAspLeuIleGlyGlySerIleLysThrValProLysProIleAlaLeu 1358
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      33414 GAAGAACCTCCTGATCTCATTGGGGGAAGTATAAAGACCGTTCCTAAGCCCATTGCCCTG 33473

Qy       1359 GluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProArgGly 1378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      33474 GAGCCCTGCTTTGGTAACAAAGCCGCAGTCCTCTCTGTATTCGTGAGGCTCCCTCGAGGA 33533

Qy       1379 LeuGlyGlyIleProProProGlyGlnSer 1388
             |||||||||||||||||||||||||||||
Db      33534 CTGGGAGGAATCCCTCCTCCTGGTCAGTCA 33563
```

```
RESULT 20
BX005139
LOCUS       BX005139              335799 bp    DNA     linear   HTG 28-FEB-2003
DEFINITION  Mus musculus chromosome X clone RP23-469F12, 3 unordered pieces.
ACCESSION   BX005139
VERSION     BX005139.4  GI:28625284
KEYWORDS    HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 335799)
  AUTHORS   Clark,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-FEB-2003) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Feb 28, 2003 this sequence version replaced gi:28460315.
            Sequence from the Mouse Genome Sequencing Consortium whole genome
            shotgun may have been used to confirm this sequence.  Sequence data
            from the whole genome shotgun alone has only been used where it has
            a phred quality of at least 30.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            -------------- Project Information
            Center project name: bM469F12
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Chemistry: Dye-terminator; 100% of reads
            Consensus quality: 165573 bases at least Q40
            Consensus quality: 165718 bases at least Q30
            Consensus quality: 165822 bases at least Q20
            Insert size: 335599; sum-of-contigs
            Insert size: 160496; 4.3% error; agarose-fp
            Quality coverage: 5.65x in Q20 bases; sum-of-contigs Quality
            coverage: 11.81x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1     5005: contig of 5005 bp in length
            *    5006     5105: gap of 100 bp
            *    5106     7812: contig of 2707 bp in length
            *    7813     7912: gap of 100 bp
            *    7913   335799: contig of 327887 bp in length.
FEATURES             Location/Qualifiers
     source          1. .335799
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="X"
                     /clone="RP23-469F12"
                     /clone_lib="RPCI-23"
     misc_feature    1. .5005
                     /note="assembly_fragment:01025
                     fragment_chain:1"
     misc_feature    5106. .7812
                     /note="assembly_fragment:00821
                     fragment_chain:1"
     misc_feature    7913. .335799
                     /note="assembly_fragment:03321"
ORIGIN

Alignment Scores:

Pred. No.:                8.77e-113      Length:          335799
Score:                    130.00         Matches:         130
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              9.10%          Indels:          0
DB:                       2              Gaps:            0

US-09-787-097-12 (1-1429) x BX005139 (1-335799)

Qy       1259 AsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAsp 1278
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63839 AATTTCACTTGGCCCATCAAAATTCAGATTGCCTTCTCCCAGCACAGCAACTTCATGGAC 63898

Qy       1279 LeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuLeuValAlaAla 1298
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63899 CTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCGCTGCTTCTGGTGGCTGCA 63958

Qy       1299 ValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeuLeuArg 1318
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63959 GTGGTCTGGAAGATCAAGCAGAGCTGTTGGGCATCCAGGCGGAGAGAGCAACTTCTTCGG 64018

Qy       1319 GluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThrAsp 1338
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64019 GAGATGCAACAGATGGCCAGCCGCCCCTTTGCTTCTGTAAACGTTGCCTTGGAAACAGAT 64078

Qy       1339 GluGluProProAspLeuIleGlyGlySerIleLysThrValProLysProIleAlaLeu 1358
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64079 GAAGAACCTCCTGATCTCATTGGGGGAAGTATAAAGACCGTTCCTAAGCCCATTGCCCTG 64138

Qy       1359 GluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProArgGly 1378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64139 GAGCCCTGCTTTGGTAACAAAGCCGCAGTCCTCTCTGTATTCGTGAGGCTCCCTCGAGGA 64198

Qy       1379 LeuGlyGlyIleProProProGlyGlnSer 1388
             |||||||||||||||||||||||||||||
Db      64199 CTGGGAGGAATCCCTCCTCCTGGTCAGTCA 64228

RESULT 21
AF301008
LOCUS       AF301008                315 bp    mRNA    linear   MAM 26-SEP-2000
DEFINITION  Bos taurus attractin mRNA, partial cds.
ACCESSION   AF301008
VERSION     AF301008.1  GI:10304438
KEYWORDS    .
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 315)
  AUTHORS   Glad,T.W., Granholm,N.H., Westby,C.A., Marshall,D.M. and Kraft,E.
  TITLE     Bovine mahogany/attractin cDNA, including exons 26-29
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 315)
  AUTHORS   Glad,T.W., Granholm,N.H., Westby,C.A., Marshall,D.M. and Kraft,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-AUG-2000) Bio/Micro, South Dakota State University,
            2140D, 252 NPB, Brookings, SD 57007, USA
FEATURES             Location/Qualifiers
     source          1. .315
                     /organism="Bos taurus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9913"
                     /tissue_type="cerebrum"
     CDS             <1. .>315
                     /note="mahogany"
                     /codon_start=2
                     /product="attractin"
                     /protein_id="AAG16254.1"
                     /db_xref="GI:10304439"
                     /translation="HSNFMDLVQFFVTFFSCFLSLLLVAAVVWKIKQSCWASRRREQL
                     LREMQQMASRPFASVNVALETDEEPPDLIGGSIKTVPKPIALEPCFGNKAAVLSVFVR
                     LPR"
```

```
ORIGIN

Alignment Scores:
Pred. No.:              3.53e-90        Length:         315
Score:                  104.00          Matches:        104
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            7.28%           Indels:         0
DB:                     4               Gaps:           0

US-09-787-097-12 (1-1429) x AF301008 (1-315)

Qy       1273 HisSerAsnPheMetAspLeuValGlnPhePheValThrPhePheSerCysPheLeuSer 1292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 CACAGCAACTTCATGGACCTGGTACAGTTCTTCGTGACCTTCTTCAGTTGTTTCCTCTCT 61

Qy       1293 LeuLeuLeuValAlaAlaValValTrpLysIleLysGlnSerCysTrpAlaSerArgArg 1312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CTGCTCCTGGTGGCTGCTGTGGTTTGGAAGATCAAACAAAGTTGTTGGGCTTCCAGACGT 121

Qy       1313 ArgGluGlnLeuLeuArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsn 1332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 AGAGAGCAACTTCTTCGAGAGATGCAGCAGATGGCCAGTCGTCCCTTTGCCTCTGTAAAT 181

Qy       1333 ValAlaLeuGluThrAspGluGluProProAspLeuIleGlyGlySerIleLysThrVal 1352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GTCGCCTTGGAAACAGATGAAGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTT 241

Qy       1353 ProLysProIleAlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPhe 1372
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CCCAAGCCCATTGCCCTGGAGCCGTGTTTTGGCAACAAAGCTGCTGTCCTCTCTGTGTTT 301

Qy       1373 ValArgLeuPro 1376
              ||||||||||||
Db        302 GTGAGGCTCCCT 313

RESULT 22
AR164813
LOCUS       AR164813                2419 bp    DNA     linear   PAT 17-OCT-2001
DEFINITION  Sequence 8 from patent US 6274339.
ACCESSION   AR164813
VERSION     AR164813.1  GI:16238022
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2419)
  AUTHORS   Moore,K. and Nagle,D.Lynn.
  TITLE     Methods and compositions for the diagnosis and treatment of body
            weight disorders, including obesity
  JOURNAL   Patent: US 6274339-A 8 14-AUG-2001;
FEATURES             Location/Qualifiers
     source          1. .2419
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:              7.67e-80        Length:         2419
Score:                  94.00           Matches:        94
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            6.58%           Indels:         0
DB:                     6               Gaps:           0

US-09-787-097-12 (1-1429) x AR164813 (1-2419)

Qy        513 AlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAsp 532
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1803 GCTCTGTACGTTCATGGTGGCTACAAGGCTTTCAGCGCCAACAAATACCGGCTTGCAGAT 1862

Qy        533 AspLeuTyrArgTyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPhe 552
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1863 GACCTCTACAGATACGATGTGGATACTCAGATGTGGACCATTCTTAAGGACAGCCGATTT 1922

Qy        553 PheArgTyrLeuHisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsn 572
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1923 TTCCGTTACTTGCATACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGAGGGAAC 1982

Qy        573 ThrHisAsnAspThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAla 592
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1983 ACACACAATGACACTTCCATGAGCCACGGTGCCAAATGCTTCTCCTCGGACTTCATGGCT 2042

Qy        593 TyrAspIleAlaCysAspArgTrpSerValLeuProArgPro 606
              ||||||||||||||||||||||||||||||||||||||
Db       2043 TATGACATTGCTTGTGACCGATGGTCAGTGCTTCCCAGACCT 2084

RESULT 23
E26747
LOCUS       E26747                  1012 bp    DNA     linear   PAT 18-JUN-2001
DEFINITION  Structure and function of novel biorhythm marker gene.
ACCESSION   E26747
VERSION     E26747.1  GI:13026325
KEYWORDS    JP 1999169185-A/1.
SOURCE      Rattus sp.
  ORGANISM  Rattus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 1012)
  AUTHORS   Mario,I., Yutaka,S., Satoru,S. and Mikio,K.
  TITLE     Structure and function of novel biorhythm marker gene
  JOURNAL   Patent: JP 1999169185-A 1 29-JUN-1999;
            AGENCY OF IND SCIENCE & TECHNOL,OTSUKA PHARMACEUT CO LTD
COMMENT     OS   Rattus sp. (Wister rat)
            PN   JP 1999169185-A/1
            PD   29-JUN-1999
            PF   12-DEC-1997 JP 1997362890
            PR
            PI   MARIO ISHIDA,YUTAKA SADAKANE,SATORU SUZUKI,MIKIO KIKUCHI PC
            C12N15/09,C07K14/47,C12Q1/68//C12N5/10,C12P21/02,(C12N15/09, PC
            C12R1:91),
            PC   (C12N5/10,C12R1:91),(C12P21/02,C12R1:91),C12N15/00,C12N5/00,
            PC   (C12N15/00,C12R1:91),(C12N5/00,C12R1:91)
            CC   Strandedness: Single;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   CDS             1. .720.
FEATURES             Location/Qualifiers
     source          1. .1012
                     /organism="Rattus sp."
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10118"
ORIGIN

Alignment Scores:
Pred. No.:              3.21e-78        Length:         1012
Score:                  92.00           Matches:        92
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            6.44%           Indels:         0
DB:                     6               Gaps:           0

US-09-787-097-12 (1-1429) x E26747 (1-1012)

Qy        849 SerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 TCCAAGCTCACTCTGACTCCATGGGTTGGTCTTCGGAAGATCAATGTATCTTACTGGTGC 372

Qy        869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        373 TGGGAGGATATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGT 432

Qy        889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        433 GATGCCGGCTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGACTGAAGGCTGCAACC 492
```

```
Qy        909 CysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSerAlaLysGln 928
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        493 TGCATCAACCCACTCAATGGAAGTGTTTGTGAAAGGCCTGCAAACCACAGTGCCAAGCAG 552

Qy        929 CysArgThrProCysAlaLeuArgThrAlaCysGly 940
              |||||||||||||||||||||||||||||||||||
Db        553 TGCCGGACACCCTGTGCCCTGCGGACAGCATGTGGT 588

RESULT 24
AB038388
LOCUS       AB038388               4500 bp    mRNA    linear   ROD 18-JAN-2001
DEFINITION  Rattus norvegicus Atrn mRNA for attractin, complete cds.
ACCESSION   AB038388
VERSION     AB038388.1  GI:12275311
KEYWORDS    attractin.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (sites)
  AUTHORS   Kuramoto,T., Kitada,K., Inui,T., Sasaki,Y., Ito,K., Hase,T.,
            Kawaguchi,S., Ogawa,Y., Nakao,K., Barsh,G.S., Nagao,M., Ushijima,T.
            and Serikawa,T.
  TITLE     Attractin/mahogany/zitter plays a critical role in myelination of
            the central nervous system
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (2), 559-564 (2001)
  MEDLINE   21143347
   PUBMED   11209055
REFERENCE   2  (bases 1 to 4500)
  AUTHORS   Kuramoto,T., Serikawa,T. and Ushijima,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-FEB-2000) Takashi Kuramoto, National Cancer Center
            Research Institute, Carcinogenesis Division; Tsukiji 5-1-1,
            Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramot@gan2.ncc.go.jp,
            Tel:81-33542-2511(ex.4521), Fax:81-35565-1753)
FEATURES             Location/Qualifiers
     source          1. .4500
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="WTC"
                     /db_xref="taxon:10116"
     gene            1. .4500
                     /gene="Atrn"
     CDS             61. .3888
                     /gene="Atrn"
                     /codon_start=1
                     /product="attractin"
                     /protein_id="BAB21018.1"
                     /db_xref="GI:12275312"
                     /translation="MVAAAAAAEATEARLRGYTTATAAPAGWKERQHRPCAATGAWRP
                     WPRAGLCLPRVLSRALSPPPLLPLLPLLFSLLLLPLPREAEAAAVAAAVSGSAAAEAK
                     ECDRPCVNGGRCNPGTGQCVCPTGWVGEQCQHCGGRFRLTGSSGFVTDGPGNYKYKTK
                     CTWLIEGQPNKIMRLRFNHFATECSWDHLYVYDGDSIYAPLIAAFSGLIVPERDGNET
                     APEVTVTSGYALLHFFSDAAYNLTGFNITYNFDMCPNNCSGRGECKSSNSSSTVECEC
                     SENWKGESCDIPHCTDNCGFPHRGICNASDTRGCSCFPHWQGPGCSIPVPANQSFWTR
                     EEYSDLKLPRASHKAEVNGNIMWVVGGYMFNHSDYSMVLAYDLASREWLSLNHSVNSV
                     VVRYGHSLALHKDKIYMYGGKIDSTGNVTNELRVFHIHNESWVLLTPKAKDQYAVVGH
                     SAHIVTLSSGRVVMLVIFGHCPLYGYISVVQEYDLEKNTWSILQTQGALVQGGYGHSS
                     VYDHRTKALYVHGGYKAFSANKYRLADDLYRYHVDTQMWTILKDSRFFRYLHTAVIVS
                     GTMLVFGGNTHNDTSMSHGAKCFSSDFMAYDIACDRWSVLPRPELHHDVNRFGHSAVL
                     HNSTMYVFGGFNSLLLSDVLVFTSEQCDAHRSEAACVAAGPGIRCLWDTQSSRCTSWE
                     LATEEQAEKLKSECFSKRTLDHDRCDQHTDCYSCTANTNDCHWCNDHCVPVNHSCTEG
                     QISIAKYDNCPKDNPMYYCNKKTSCRSCALDQNCQWEPRNQECIALPENICGIGWHLV
                     GNSCLKITTAKENYDNAKLSCRNHNAPLASLTSQKKVEFVLKQLRLMQSSQSTSKLTL
                     TPWVGLRKINVSYWCWEDMSPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCIN
                     PLNGSVCERPANHSAKQCRTPCALRTACGECTSSSSECMWCSNMKQCVDSNAYVASFP
                     FGQCMEWYTMSSCPPENCSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKM
                     PSHASTGNVYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCEDLTT
                     GKHCETCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGEECQLCEVE
                     NRYQGNPLKGTCYYTLLIDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASK
                     NFNLNITWATSFPAGTQTGEEVPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSN
                     FTWPIKIQVRVTS"
ORIGIN

Alignment Scores:
Pred. No.:                1.06e-77        Length:          4500
Score:                    92.00           Matches:         353
Percent Similarity:       95.15%          Conservative:    0
Best Local Similarity:    95.15%          Mismatches:      9
Query Match:              6.44%           Indels:          18
DB:                       10              Gaps:            0

US-09-787-097-12 (1-1429) x AB038388 (1-4500)

Qy         849 SerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2614 TCCAAGCTCACTCTGACTCCATGGGTTGGTCTTCGGAAGATCAATGTATCTTACTGGTGC 2673

Qy         869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2674 TGGGAGGATATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGT 2733

Qy         889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2734 GATGCCGGCTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGACTGAAGGCTGCAACC 2793

Qy         909 CysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSerAlaLysGln 928
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2794 TGCATCAACCCACTCAATGGAAGTGTTTGTGAAAGGCCTGCAAACCACAGTGCCAAGCAG 2853

Qy         929 CysArgThrProCysAlaLeuArgThrAlaCysGlyAsp-CysThrSerGly-SerSerG 948
               |||||||||||||||||||||||||||||||||||||||  ||||||||  |||||||
Db        2854 TGCCGGACACCCTGTGCCCTGCGGACAGCATGTGGTGA-GTGCACCAGCAG-CAGCTCCG 2911

Qy         948 luCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrValAlaSerP 968
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2912 AATGCATGTGGTGCAGCAACATGAAGCAGTGCGTAGACTCCAATGCCTACGTAGCCTCCT 2971

Qy         968 heProPheGlyGlnCysMetGluTrpTyrThrMetSerThr-CysProProGluAsnCys 987
               |||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||
Db        2972 TTCCTTTCGGCCAGTGTATGGAGTGGTATACAATGAGCAG-CTGCCCACCTGAAAATTGC 3030

Qy         988 SerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCysThrAsp 1007
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3031 TCTGGCTACTGTACCTGCAGCCATTGCTTGGAGCAGCCAGGCTGTGGTTGGTGTACTGAT 3090

Qy        1008 ProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyProValLysMet 1027
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3091 CCTAGCAATACTGGGAAAGGCAAATGTATTGAGGGAAGCTATAAAGGACCCGTGAAGATG 3150

Qy        1028 ProSerGlnAlaPro--ThrGlyAsnPhe-TyrProGlnProLeuLeuAsnSerSerMet 1046
               ||||||   |||  |||||||||  |||||||||||||||||||||||||||||||||||
Db        3151 CCTTCACA-TG-CCTCTACAGGAAATGT-GTACCCACAGCCCCTTCTGAACTCCAGCATG 3207

Qy        1047 CysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnCysAsn 1066
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3208 TGTCTAGAGGACAGCAGATACAACTGGTCTTTTATTCACTGTCCAGCTTGCCAGTGCAAT 3267

Qy        1067 GlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsn-LeuThrThrGl 1086
               ||||||||||||||||||||||||||||||||||||||||||||||  ||||||||||
Db        3268 GGACACAGCAAGTGCATCAACCAGAGCATCTGTGAGAAGTGTGAGGA-CCTGACCACAGG 3326

Qy        1086 yLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyGlyLysCy 1106
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3327 CAAGCACTGCGAGACCTGCATATCTGGCTTCTATGGTGACCCGACTAATGGAGGCAAATG 3386

Qy        1106 sGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysPh 1126
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3387 TCAGCCATGCAAGTGCAACGGGCACGCATCACTGTGCAACACCAACACTGGCAAGTGCTT 3446

Qy        1126 eCysThrThrLysGlyValLysGlyAsp-GluCysGlnLeuCysGluValGluAsnArgT 1146
               ||||||||||||||||||||||||  |||||||||||||||||||||||||||||||||
Db        3447 CTGTACCACCAAGGGCGTCAAGGGAGA-GGAGTGCCAGCTATGTGAGGTAGAAAATCGAT 3505
```

```
Qy       1146 yrGlnGlyAsnProLeuArg-GlyThrCysTyrTyrThrLeuLeuIleAspTyrGlnPhe 1165
             ||||||||||||||||||    ||||||||||||||||||||||||||||||||||||||
Db       3506 ACCAGGGAAACCCTCTCAA-AGGAACATGTTACTATACCCTTCTCATTGACTATCAGTTC 3564

Qy       1166 ThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIleAsnPheValAlaThr 1185
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3565 ACTTTTAGCCTGTCCCAGGAAGACGACCGCTATTACACAGCCATCAATTTTGTGGCTACT 3624

Qy       1186 ProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeu 1205
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3625 CCTGATGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAAAACTTCAACCTC 3684

Qy       1206 AsnIleThrTrpAla 1210
             ||||||||||||||||
Db       3685 AACATCACCTGGGCC 3699

RESULT 25
AR273223/c
LOCUS       AR273223                246 bp    DNA     linear   PAT 10-APR-2003
DEFINITION  Sequence 966 from patent US 6504010.
ACCESSION   AR273223
VERSION     AR273223.1  GI:29705108
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 246)
  AUTHORS   Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S.,
            Carter,D., Retter,M.W., Mannion,J. and Fan,L.
  TITLE     Compositions and methods for the therapy and diagnosis of lung
            cancer
  JOURNAL   Patent: US 6504010-A 966 07-JAN-2003;
FEATURES             Location/Qualifiers
     source          1. .246
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:              4.38e-69        Length:         246
Score:                  82.00           Matches:        82
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            5.74%           Indels:         0
DB:                     6               Gaps:           0

US-09-787-097-12 (1-1429) x AR273223 (1-246)

Qy        839 LeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGly 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy        859 LeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSer 878
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy        879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGlu 898
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67

Qy        899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy        919 GluArg 920
              ||||||
Db          6 GAAAGG 1

RESULT 26
AR276804/c
LOCUS       AR276804                246 bp    DNA     linear   PAT 10-APR-2003
DEFINITION  Sequence 966 from patent US 6509448.
ACCESSION   AR276804
VERSION     AR276804.1  GI:29710451
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 246)
  AUTHORS   Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S.,
            Carter,D., Retter,M.W., Mannion,J., Fan,L. and Wang,A.
  TITLE     Compositions and methods for the therapy and diagnosis of lung
            cancer
  JOURNAL   Patent: US 6509448-A 966 21-JAN-2003;
FEATURES             Location/Qualifiers
     source          1. .246
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:              4.38e-69        Length:         246
Score:                  82.00           Matches:        82
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            5.74%           Indels:         0
DB:                     6               Gaps:           0

US-09-787-097-12 (1-1429) x AR276804 (1-246)

Qy        839 LeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGly 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy        859 LeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSer 878
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy        879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGlu 898
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67

Qy        899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy        919 GluArg 920
              ||||||
Db          6 GAAAGG 1

RESULT 27
AR407079/c
LOCUS       AR407079                246 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 966 from patent US 6630574.
ACCESSION   AR407079
VERSION     AR407079.1  GI:40156890
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 246)
  AUTHORS   Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S.,
            Carter,D., Retter,M.W., Mannion,J. and Fan,L.
  TITLE     Compositions and methods for the therapy and diagnosis of lung
            cancer
  JOURNAL   Patent: US 6630574-A 966 07-OCT-2003;
FEATURES             Location/Qualifiers
     source          1. .246
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:              4.38e-69        Length:         246
Score:                  82.00           Matches:        82
Percent Similarity:     100.00%         Conservative:   0
```

```
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            5.74%          Indels:        0
DB:                     6              Gaps:          0

US-09-787-097-12 (1-1429) x AR407079 (1-246)

Qy        839 LeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGly 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy        859 LeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSer 878
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy        879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGlu 898
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67

Qy        899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy        919 GluArg 920
              ||||||
Db          6 GAAAGG 1

RESULT 28
AX368256/c
LOCUS       AX368256                246 bp    DNA     linear   PAT 16-FEB-2002
DEFINITION  Sequence 966 from Patent WO0204514.
ACCESSION   AX368256
VERSION     AX368256.1  GI:18856329
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Wang,T., Watanabe,Y., Henderson,R.A., Johnson,J.C., Retter,M.W.,
            Marnerakis,M., Carter,D., Fanger,G.R., Vedvick,T.S., Bangur,C.S.,
            Mcnabb,A., Fanger,N., Switzer,A., Mcneill,P.D. and Clapper,J.D.
  TITLE     Compositions and methods for the therapy and diagnosis of lung
            cancer
  JOURNAL   Patent: WO 0204514-A 966 17-JAN-2002;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .246
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              4.38e-69       Length:        246
Score:                  82.00          Matches:       82
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            5.74%          Indels:        0
DB:                     6              Gaps:          0

US-09-787-097-12 (1-1429) x AX368256 (1-246)

Qy        839 LeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGly 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy        859 LeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSer 878
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy        879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGlu 898
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67

Qy        899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy        919 GluArg 920
              ||||||
Db          6 GAAAGG 1

RESULT 29
AL954766_2
WPCOMMENT
  Sequence split into 4 fragments  LOCUS AL954766 Accession AL954766
     Fragment Name      Begin      End
     AL954766_0             1    110000
     AL954766_1        100001    210000
     AL954766_2        200001    310000
     AL954766_3        300001    356753
Continuation (3 of 4) of AL954766 from base 200001 (AL954766 Mus musculus chromosome X c

Alignment Scores:
Pred. No.:              4.51e-64       Length:        110000
Score:                  79.00          Matches:       79
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            5.53%          Indels:        0
DB:                     2              Gaps:          0

US-09-787-097-12 (1-1429) x AL954766_2 (1-110000)

Qy       1259 AsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAsp 1278
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      60740 AATTTCACTTGGCCCATCAAAATTCAGATTGCCTTCTCCCAGCACAGCAACTTCATGGAC 60799

Qy       1279 LeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuLeuValAlaAla 1298
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      60800 CTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCGCTGCTTCTGGTGGCTGCA 60859

Qy       1299 ValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeuLeuArg 1318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      60860 GTGGTCTGGAAGATCAAGCAGAGCTGTTGGGCATCCAGGCGGAGAGAGCAACTTCTTCGG 60919

Qy       1319 GluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThr 1337
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      60920 GAGATGCAACAGATGGCCAGCCGCCCCTTTGCTTCTGTAAACGTTGCCTTGGAAACG 60976

RESULT 30
BX004762_3
WPCOMMENT
  Sequence split into 5 fragments  LOCUS BX004762 Accession BX004762
     Fragment Name      Begin      End
     BX004762_0             1    110000
     BX004762_1        100001    210000
     BX004762_2        200001    310000
     BX004762_3        300001    410000
     BX004762_4        400001    442165
Continuation (4 of 5) of BX004762 from base 300001 (BX004762 Mus musculus chromosome X c

Alignment Scores:
Pred. No.:              4.51e-64       Length:        110000
Score:                  79.00          Matches:       79
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            5.53%          Indels:        0
DB:                     2              Gaps:          0

US-09-787-097-12 (1-1429) x BX004762_3 (1-110000)

Qy       1259 AsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAsp 1278
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      46152 AATTTCACTTGGCCCATCAAAATTCAGATTGCCTTCTCCCAGCACAGCAACTTCATGGAC 46211

Qy       1279 LeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuLeuValAlaAla 1298
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      46212 CTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCGCTGCTTCTGGTGGCTGCA 46271

Qy       1299 ValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeuLeuArg 1318
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      46272 GTGGTCTGGAAGATCAAGCAGAGCTGTTGGGCATCCAGGCGGAGAGAGCAACTTCTTCGG 46331

Qy       1319 GluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThr 1337
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      46332 GAGATGCAACAGATGGCCAGCCGCCCCTTTGCTTCTGTAAACGTTGCCTTGGAAACG 46388

RESULT 31
H009384S15
LOCUS       H009384S15              3257 bp    DNA     linear   PRI 14-DEC-2000
DEFINITION  Homo sapiens attractin precursor (ATRN) gene, exons 17 and 18.
ACCESSION   AF218903
VERSION     AF218903.1  GI:8118069
KEYWORDS    .
SEGMENT     15 of 27
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3257)
  AUTHORS   Tang,W., Gunn,T.M., McLaughlin,D.F., Barsh,G.S., Schlossman,S.F.
            and Duke-Cohan,J.S.
  TITLE     Secreted and membrane attractin result from alternative splicing of
            the human ATRN gene
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
  MEDLINE   20283944
   PUBMED   10811918
REFERENCE   2  (bases 1 to 3257)
  AUTHORS   Tang,W., Gunn,T.M., McLaughlin,D.F., Freeman,G.J., Barsh,G.S.,
            Schlossman,S.F. and Duke-Cohan,J.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
            Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
            USA
FEATURES             Location/Qualifiers
     source          1. .3257
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /map="20p13"
     exon            1441. .1626
                     /gene="ATRN"
                     /number=17
     exon            2189. .2422
                     /gene="ATRN"
                     /number=18
ORIGIN

Alignment Scores:
Pred. No.:              2.26e-63        Length:         3257
Score:                  77.00           Matches:        77
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            5.39%           Indels:         0
DB:                     9               Gaps:           0

US-09-787-097-12 (1-1429) x H009384S15 (1-3257)

Qy        985 GluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrp 1004
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2191 GAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGG 2250

Qy       1005 CysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyPro 1024
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2251 TGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCA 2310

Qy       1025 ValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSer 1044
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2311 GTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCC 2370

Qy       1045 SerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysPro 1061
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2371 AGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCA 2421

RESULT 32
HSDJ741H3
LOCUS       HSDJ741H3              104907 bp    DNA     linear   PRI 23-FEB-2001
DEFINITION  Human DNA sequence from clone RP4-741H3 on chromosome 20. 'Contains
            the 3' end of the ATRN gene for attractin (with dipeptidylpeptidase
            IV activity) (KIAA0548), ESTs, STSs and GSSs, complete sequence.
ACCESSION   AL132773
VERSION     AL132773.14  GI:7159777
KEYWORDS    HTG; ATRN; attractin; KIAA0548.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 104907)
  AUTHORS   Sehra,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Mar 6, 2000 this sequence version replaced gi:7007301.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 20, constructed by the Sanger Centre Chromosome 20
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr20
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. RP4-741H3 is from the
            library RPCI-4 constructed by the group of Pieter de Jong. For
            further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pCYPAC2
            This sequence is the entire insert of clone RP4-741H3 The true
            right end of clone RP4-581P3 is at 79521 in this sequence.
FEATURES             Location/Qualifiers
     source          1. .104907
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /clone="RP4-741H3"
                     /clone_lib="RPCI-4"
     gene            2. .103785
                     /gene="ATRN"
     mRNA            join(<2. .161,1842. .2010,12065. .12155,13334. .13577,
                     14948. .15131,15881. .16035,18020. .18104,25403. .25623,
                     28499. .28620,29531. .29669,31254. .31444,34807. .35026,
                     36570. .36755,37319. .37552,43841. .43978,47151. .47244,
                     49049. .49101,50578. .50672,53618. .53696,56777. .56934,
                     77183. .77244,86983. .87061,91500. .91607,96826. .96940,
                     99395. .103785)
                     /gene="ATRN"
                     /product="dJ741H3.1.2 (attractin (with dipeptidylpeptidase
```

IV activity) membrane bound isoform (KIAA0548))"
/note="match: cDNAs: Em:AF116897 Em:AF119821 Em:AF106861
Em:AF034957 Em:AB011106 Em:AK000356 Em:AB011120
match: ESTs: Em:R87660 Em:H11351 Em:AA350292 Em:T69367
Em:AA368140 Em:AI372811 Em:AW141339 Em:AL048842"
/evidence=not_experimental
CDS join(<2. .161,1842. .2010,12065. .12155,13334. .13577,
14948. .15131,15881. .16035,18020. .18104,25403. .25623,
28499. .28620,29531. .29669,31254. .31444,34807. .35026,
36570. .36755,37319. .37552,43841. .43978,47151. .47244,
49049. .49101,50578. .50672,53618. .53696,56777. .56934,
77183. .77244,86983. .87061,91500. .91607,96826. .96940,
99395. .99519)
/gene="ATRN"
/note="match: proteins: Tr:O60283 Tr:Q9WU60 Tr:Q9WU77
Tr:O60295 Tr:O95414 Tr:O75882 Sw:Q19981 Sw:P19137
Sw:P25391"
/codon_start=1
/evidence=not_experimental
/product="dJ741H3.1.2 (attractin (with dipeptidylpeptidase
IV activity) membrane bound isoform (KIAA0548))"
/protein_id="CAC32457.1"
/db_xref="GI:13160052"
/db_xref="GOA:O75882"
/db_xref="SWISS-PROT:O75882"
/translation="ISNSSDTVECECSENWKGEACDIPHCTDNCGFPHRGICNSSDVR
GCSCFSDWQGPGCSVPVPANQSFWTREEYSNLKLPRASHKAVVNGNIMWVVGGYMFNH
SDYNMVLAYDLASREWLPLNRSVNNVVVRYGHSLALYKDKIYMYGGKIDSTGNVTNEL
RVFHIHNESWVLLTPKAKEQYAVVGHSAHIVTLKNGRVVMLVIFGHCPLYGYISNVQE
YDLDKNTWSILHTQGALVQGGYGHSSVYDHRTRALYVHGGYKAFSANKYRLADDLYRY
DVDTQMWTILKDSRFFRYLHTAVIVSGTMLVFGGNTHNDTSMSHGAKCFSSDFMAYDI
ACDRWSVLPRPDLHHDVNRFGHSAVLHNSTMYVFGGFNSLLLSDILVFTSEQCDAHRS
EAACLAAGPGIRCVWNTGSSQCISWALATDEQEEKLKSECFSKRTLDHDRCDQHTDCY
SCTANTNDCHWCNDHCVPRNHSCSEGQISIFRYENCPKDNPMYYCNKKTSCRSCALDQ
NCQWEPRNQECIALPENICGIGWHLVGNSCLKITTAKENYDNAKLFCRNHNALLASLT
TQKKVEFVLKQLRIMQSSQSMSKLTLTPWVGLRKINVSYWCWEDMSPFTNSLLQWMPS
EPSDAGFCGILSEPSTRGLKAATCINPLNGSVCERPANHSAKQCRTPCALRTACGDCT
SGSSECMWCSNMKQCVDSNAYVASFPFGQCMEWYTMSTCPPENCSGYCTCSHCLEQPG
CGWCTDPSNTGKGKCIEGSYKGPVKMPSQAPTGNFYPQPLLNSSMCLEDSRYNWSFIH
CPACQCNGHSKCINQSICEKCENLTTGKHCETCISGFYGDPTNGGKCQPCKCNGHASL
CNTNTGKCFCTTKGVKGDECQLCEVENRYQGNPLRGTCYYTLLIDYQFTFSLSQEDDR
YYTAINFVATPDEQNRDLDMFINASKNFNLNITWAASFSAGTQAGEEMPVVSKTNIKE
YKDSFSNEKFDFRNHPNITFFVYVSNFTWPIKIQVQTEQ"
CDS join(<2. .161,1842. .2010,12065. .12155,13334. .13577,
14948. .15131,15881. .16035,18020. .18104,25403. .25623,
28499. .28620,29531. .29669,31254. .31444,34807. .35026,
36570. .36755,37319. .37552,43841. .43978,47151. .47244,
49049. .49101,50578. .50672,53618. .53696,56777. .56934,
68500. .68517)
/gene="ATRN"
/note="match: proteins: Tr:O75882 Tr:Q9NZ58 Tr:O95414"
/codon_start=1
/evidence=not_experimental
/product="dJ741H3.1.1 (attractin (with dipeptidylpeptidase
IV activity) secreted isoform)"
/protein_id="CAC32456.1"
/db_xref="GI:13160051"
/db_xref="GOA:O75882"
/db_xref="SWISS-PROT:O75882"
/translation="ISNSSDTVECECSENWKGEACDIPHCTDNCGFPHRGICNSSDVR
GCSCFSDWQGPGCSVPVPANQSFWTREEYSNLKLPRASHKAVVNGNIMWVVGGYMFNH
SDYNMVLAYDLASREWLPLNRSVNNVVVRYGHSLALYKDKIYMYGGKIDSTGNVTNEL
RVFHIHNESWVLLTPKAKEQYAVVGHSAHIVTLKNGRVVMLVIFGHCPLYGYISNVQE
YDLDKNTWSILHTQGALVQGGYGHSSVYDHRTRALYVHGGYKAFSANKYRLADDLYRY
DVDTQMWTILKDSRFFRYLHTAVIVSGTMLVFGGNTHNDTSMSHGAKCFSSDFMAYDI
ACDRWSVLPRPDLHHDVNRFGHSAVLHNSTMYVFGGFNSLLLSDILVFTSEQCDAHRS
EAACLAAGPGIRCVWNTGSSQCISWALATDEQEEKLKSECFSKRTLDHDRCDQHTDCY
SCTANTNDCHWCNDHCVPRNHSCSEGQISIFRYENCPKDNPMYYCNKKTSCRSCALDQ
NCQWEPRNQECIALPENICGIGWHLVGNSCLKITTAKENYDNAKLFCRNHNALLASLT
TQKKVEFVLKQLRIMQSSQSMSKLTLTPWVGLRKINVSYWCWEDMSPFTNSLLQWMPS
EPSDAGFCGILSEPSTRGLKAATCINPLNGSVCERPANHSAKQCRTPCALRTACGDCT
SGSSECMWCSNMKQCVDSNAYVASFPFGQCMEWYTMSTCPPENCSGYCTCSHCLEQPG
CGWCTDPSNTGKGKCIEGSYKGPVKMPSQAPTGNFYPQPLLNSSMCLEDSRYNWSFIH
CPACQCNGHSKCINQSICEKCENLTTGKHCETCISGFYGDPTNGGKCQPCKCNGHASL
CNTNTGKCFCTTKGVKGDECQLCEVENRYQGNPLRGTCYYTLLIDYQFTFSLSQEDDR
YYTAINFVATPDEQNRDLDMFINASKNFNLNITWAASFSAGTQAGEEMPVVSKTNIKE
YKDSFSNEKFDFRNHPNITFFVYVSNFTWPIKIQVQTEQ"
misc_feature 17. .569
/gene="ATRN"
/note="match: STS: Em:AL110053 Em:HS741H3S"
repeat_region 903. .990
/note="MER2 repeat: matches 257. .337 of consensus"
repeat_region 991. .1297
/note="AluSx repeat: matches 1. .307 of consensus"
repeat_region 1298. .1520
/note="MER2 repeat: matches 1. .257 of consensus"
repeat_region 2718. .2942
/note="AluSq repeat: matches 1. .225 of consensus"
repeat_region 2943. .3127
/note="L1P repeat: matches 4678. .4858 of consensus"
misc_feature complement(3203. .3804)
/note="match: GSS: Em:AQ381369"
repeat_region 3560. .3672
/note="L2 repeat: matches 216. .330 of consensus"
repeat_region 3881. .3977
/note="MER81 repeat: matches 2. .110 of consensus"
repeat_region 4326. .4513
/note="L2 repeat: matches 738. .924 of consensus"
repeat_region 4539. .4682
/note="MER3 repeat: matches 1. .149 of consensus"
repeat_region 4683. .5195
/note="L2 repeat: matches 1383. .1936 of consensus"
repeat_region 5209. .5762
/note="MLT1E repeat: matches 2. .562 of consensus"
repeat_region 5770. .5835
/note="L2 repeat: matches 1949. .2024 of consensus"
repeat_region 6074. .6708
/note="L2 repeat: matches 2025. .2732 of consensus"
repeat_region 6998. .7305
/note="AluY repeat: matches 1. .309 of consensus"
repeat_region 7663. .7974
/note="AluY repeat: matches 1. .311 of consensus"
misc_feature complement(9805. .10470)
/note="match: GSS: Em:AQ484998"
misc_feature 9825. .10261
/gene="ATRN"
/note="match: GSS: Em:AQ338818"
repeat_region 9963. .10167
/note="HAL1 repeat: matches 618. .819 of consensus"
misc_feature 10173. .10450
/gene="ATRN"
/note="match: GSS: Em:AQ413104"
repeat_region 10902. .11155
/note="L1ME3 repeat: matches 5637. .5902 of consensus"
repeat_region 12706. .12802
/note="L1MB2 repeat: matches 6075. .6168 of consensus"
repeat_region 12803. .13235
/note="L1MA4A repeat: matches 5584. .6168 of consensus"
repeat_region 13713. .14023
/note="AluSx repeat: matches 3. .311 of consensus"
repeat_region 14407. .14697
/note="AluJo repeat: matches 1. .293 of consensus"

Alignment Scores:

Pred. No.:	3.65e-62	Length:	104907
Score:	77.00	Matches:	77
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.39%	Indels:	0
DB:	9	Gaps:	0

US-09-787-097-12 (1-1429) x HSDJ741H3 (1-104907)

```
Qy        985 GluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrp 1004
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37321 GAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGG 37380

Qy       1005 CysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyPro 1024
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37381 TGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCA 37440

Qy       1025 ValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSer 1044
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37441 GTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCC 37500

Qy       1045 SerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysPro 1061
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37501 AGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCA 37551

RESULT 33
AC013324/c
LOCUS       AC013324              118540 bp    DNA     linear   HTG 13-JUL-2000
DEFINITION  Homo sapiens chromosome 2 clone RP11-388K24 map 2, LOW-PASS
            SEQUENCE SAMPLING.
ACCESSION   AC013324
VERSION     AC013324.3  GI:9123914
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 118540)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 2, clone RP11-388K24
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 118540)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baldwin,J., Barna,N., Beckerly,R., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
            Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
            Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
            McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
            Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
            Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-NOV-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Jul 13, 2000 this sequence version replaced gi:6289165.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L3787
                Center clone name: 388_K_24
            --------------.
            * NOTE: This record contains 125 individual
            * sequencing reads that have not been assembled into
            * contigs. Runs of N are used to separate the reads
            * and the order in which they appear is completely
            * arbitrary. Low-pass sequence sampling is useful for
            * identifying clones that may be gene-rich and allows
            * overlap relationships among clones to be deduced.
            * However, it should not be assumed that this clone
            * will be sequenced to completion. In the event that
            * the record is updated, the accession number will
            * be preserved.
            *       1      856: contig of 856 bp in length
            *     857      956: gap of 100 bp
            *     957     1821: contig of 865 bp in length
            *    1822     1921: gap of 100 bp
            *    1922     2776: contig of 855 bp in length
            *    2777     2876: gap of 100 bp
            *    2877     3767: contig of 891 bp in length
            *    3768     3867: gap of 100 bp
            *    3868     4709: contig of 842 bp in length
            *    4710     4809: gap of 100 bp
            *    4810     5655: contig of 846 bp in length
            *    5656     5755: gap of 100 bp
            *    5756     6627: contig of 872 bp in length
            *    6628     6727: gap of 100 bp
            *    6728     7586: contig of 859 bp in length
            *    7587     7686: gap of 100 bp
            *    7687     8541: contig of 855 bp in length
            *    8542     8641: gap of 100 bp
            *    8642     9524: contig of 883 bp in length
            *    9525     9624: gap of 100 bp
            *    9625    10509: contig of 885 bp in length
            *   10510    10609: gap of 100 bp
            *   10610    11474: contig of 865 bp in length
            *   11475    11574: gap of 100 bp
            *   11575    12441: contig of 867 bp in length
            *   12442    12541: gap of 100 bp
            *   12542    13409: contig of 868 bp in length
            *   13410    13509: gap of 100 bp
            *   13510    14439: contig of 930 bp in length
            *   14440    14539: gap of 100 bp
            *   14540    15433: contig of 894 bp in length
            *   15434    15533: gap of 100 bp
            *   15534    16442: contig of 909 bp in length
            *   16443    16542: gap of 100 bp
            *   16543    17421: contig of 879 bp in length
            *   17422    17521: gap of 100 bp
            *   17522    18380: contig of 859 bp in length
            *   18381    18480: gap of 100 bp
            *   18481    19345: contig of 865 bp in length
            *   19346    19445: gap of 100 bp
            *   19446    20327: contig of 882 bp in length
            *   20328    20427: gap of 100 bp
            *   20428    21297: contig of 870 bp in length
            *   21298    21397: gap of 100 bp
            *   21398    22055: contig of 658 bp in length
            *   22056    22155: gap of 100 bp
            *   22156    23005: contig of 850 bp in length
            *   23006    23105: gap of 100 bp
            *   23106    24016: contig of 911 bp in length
            *   24017    24116: gap of 100 bp
            *   24117    24972: contig of 856 bp in length
            *   24973    25072: gap of 100 bp
            *   25073    25977: contig of 905 bp in length
            *   25978    26077: gap of 100 bp
            *   26078    26952: contig of 875 bp in length
            *   26953    27052: gap of 100 bp
            *   27053    27911: contig of 859 bp in length
            *   27912    28011: gap of 100 bp
            *   28012    28868: contig of 857 bp in length
            *   28869    28968: gap of 100 bp
            *   28969    29624: contig of 656 bp in length
            *   29625    29724: gap of 100 bp
            *   29725    30574: contig of 850 bp in length
            *   30575    30674: gap of 100 bp
            *   30675    31530: contig of 856 bp in length
            *   31531    31630: gap of 100 bp
            *   31631    32491: contig of 861 bp in length
            *   32492    32591: gap of 100 bp
            *   32592    33467: contig of 876 bp in length
            *   33468    33567: gap of 100 bp
            *   33568    34437: contig of 870 bp in length
            *   34438    34537: gap of 100 bp
```

```
*    34538    35404: contig of 867 bp in length
*    35405    35504: gap of 100 bp
*    35505    36368: contig of 864 bp in length
*    36369    36468: gap of 100 bp
*    36469    37353: contig of 885 bp in length
*    37354    37453: gap of 100 bp
*    37454    38310: contig of 857 bp in length
*    38311    38410: gap of 100 bp
*    38411    39267: contig of 857 bp in length
*    39268    39367: gap of 100 bp
*    39368    40217: contig of 850 bp in length
*    40218    40317: gap of 100 bp
*    40318    41226: contig of 909 bp in length
*    41227    41326: gap of 100 bp
*    41327    42195: contig of 869 bp in length
*    42196    42295: gap of 100 bp
*    42296    43180: contig of 885 bp in length
*    43181    43280: gap of 100 bp
*    43281    44143: contig of 863 bp in length
*    44144    44243: gap of 100 bp
*    44244    45131: contig of 888 bp in length
*    45132    45231: gap of 100 bp
*    45232    46149: contig of 918 bp in length
*    46150    46249: gap of 100 bp
*    46250    47122: contig of 873 bp in length
*    47123    47222: gap of 100 bp
*    47223    48101: contig of 879 bp in length
*    48102    48201: gap of 100 bp
*    48202    49041: contig of 840 bp in length
*    49042    49141: gap of 100 bp
*    49142    50059: contig of 918 bp in length
*    50060    50159: gap of 100 bp
*    50160    51029: contig of 870 bp in length
*    51030    51129: gap of 100 bp
*    51130    51995: contig of 866 bp in length
*    51996    52095: gap of 100 bp
*    52096    52971: contig of 876 bp in length
*    52972    53071: gap of 100 bp
*    53072    53942: contig of 871 bp in length
*    53943    54042: gap of 100 bp
*    54043    54923: contig of 881 bp in length
*    54924    55023: gap of 100 bp
*    55024    55877: contig of 854 bp in length
*    55878    55977: gap of 100 bp
*    55978    56851: contig of 874 bp in length
*    56852    56951: gap of 100 bp
*    56952    57808: contig of 857 bp in length
*    57809    57908: gap of 100 bp
*    57909    58755: contig of 847 bp in length
*    58756    58855: gap of 100 bp
*    58856    59709: contig of 854 bp in length
*    59710    59809: gap of 100 bp
*    59810    60521: contig of 712 bp in length
*    60522    60621: gap of 100 bp
*    60622    61501: contig of 880 bp in length
*    61502    61601: gap of 100 bp
*    61602    62490: contig of 889 bp in length
*    62491    62590: gap of 100 bp
*    62591    63450: contig of 860 bp in length
*    63451    63550: gap of 100 bp
*    63551    64094: contig of 544 bp in length
*    64095    64194: gap of 100 bp
*    64195    65022: contig of 828 bp in length
*    65023    65122: gap of 100 bp
*    65123    65983: contig of 861 bp in length
*    65984    66083: gap of 100 bp
*    66084    66922: contig of 839 bp in length
*    66923    67022: gap of 100 bp
*    67023    67908: contig of 886 bp in length
*    67909    68008: gap of 100 bp
*    68009    68879: contig of 871 bp in length
*    68880    68979: gap of 100 bp
```

```
Alignment Scores:
Pred. No.:                4.03e-62        Length:          118540
Score:                    77.00           Matches:         77
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              5.39%           Indels:          0
DB:                       2               Gaps:            0

US-09-787-097-12 (1-1429) x AC013324 (1-118540)

Qy        985 GluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrp 1004
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      31009 GAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGG 30950

Qy       1005 CysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyPro 1024
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      30949 TGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCA 30890

Qy       1025 ValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSer 1044
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      30889 GTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCC 30830

Qy       1045 SerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysPro 1061
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      30829 AGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCA 30779

RESULT 34
AC055771/c
LOCUS       AC055771               167357 bp    DNA     linear   HTG 24-AUG-2002
DEFINITION  Homo sapiens chromosome 20 clone RP11-561P16 map 20, WORKING DRAFT
            SEQUENCE, 21 unordered pieces.
ACCESSION   AC055771
VERSION     AC055771.2  GI:7770449
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 167357)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 20, clone RP11-561P16
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 167357)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-APR-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 167357)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
```

```
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On May 12, 2000 this sequence version replaced gi:7582593.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L5007
                Center clone name: 561_P_16
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 156504 bases at least Q40
                Consensus quality: 161679 bases at least Q30
                Consensus quality: 163964 bases at least Q20
                Insert size: 170000; agarose-fp
                Insert size: 165357; sum-of-contigs
                Quality coverage: 4.5 in Q20 bases; agarose-fp
                Quality coverage: 4.6 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 21 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1216: contig of 1216 bp in length
            *     1217     1316: gap of 100 bp
            *     1317     2985: contig of 1669 bp in length
            *     2986     3085: gap of 100 bp
            *     3086     4628: contig of 1543 bp in length
            *     4629     4728: gap of 100 bp
            *     4729     7507: contig of 2779 bp in length
            *     7508     7607: gap of 100 bp
            *     7608    11292: contig of 3685 bp in length
            *    11293    11392: gap of 100 bp
            *    11393    14213: contig of 2821 bp in length
            *    14214    14313: gap of 100 bp
            *    14314    17828: contig of 3515 bp in length
            *    17829    17928: gap of 100 bp
            *    17929    21068: contig of 3140 bp in length
            *    21069    21168: gap of 100 bp
            *    21169    25035: contig of 3867 bp in length
            *    25036    25135: gap of 100 bp
            *    25136    28331: contig of 3196 bp in length
            *    28332    28431: gap of 100 bp
            *    28432    33595: contig of 5164 bp in length
            *    33596    33695: gap of 100 bp
            *    33696    40346: contig of 6651 bp in length
            *    40347    40446: gap of 100 bp
            *    40447    43807: contig of 3361 bp in length
            *    43808    43907: gap of 100 bp
            *    43908    48727: contig of 4820 bp in length
            *    48728    48827: gap of 100 bp
            *    48828    56083: contig of 7256 bp in length
            *    56084    56183: gap of 100 bp
            *    56184    67065: contig of 10882 bp in length
            *    67066    67165: gap of 100 bp
            *    67166    82881: contig of 15716 bp in length
            *    82882    82981: gap of 100 bp
            *    82982   101845: contig of 18864 bp in length
            *   101846   101945: gap of 100 bp
            *   101946   122755: contig of 20810 bp in length
            *   122756   122855: gap of 100 bp
            *   122856   142457: contig of 19602 bp in length
            *   142458   142557: gap of 100 bp
            *   142558   167357: contig of 24800 bp in length.
FEATURES             Location/Qualifiers
     source          1. .167357
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /map="20"
                     /clone="RP11-561P16"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1216
                     /note="assembly_fragment"
     misc_feature    1317. .2985
                     /note="assembly_fragment"
     misc_feature    3086. .4628
                     /note="assembly_fragment"
     misc_feature    4729. .7507
                     /note="assembly_fragment"
     misc_feature    7608. .11292
                     /note="assembly_fragment"
     misc_feature    11393. .14213
                     /note="assembly_fragment"
     misc_feature    14314. .17828
                     /note="assembly_fragment"
     misc_feature    17929. .21068
                     /note="assembly_fragment"
     misc_feature    21169. .25035
                     /note="assembly_fragment"
     misc_feature    25136. .28331
                     /note="assembly_fragment"
     misc_feature    28432. .33595
                     /note="assembly_fragment"
     misc_feature    33696. .40346
                     /note="assembly_fragment"
     misc_feature    40447. .43807
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:right"
     misc_feature    43908. .48727
                     /note="assembly_fragment"
     misc_feature    48828. .56083
                     /note="assembly_fragment"
     misc_feature    56184. .67065
                     /note="assembly_fragment"
     misc_feature    67166. .82881
                     /note="assembly_fragment"
     misc_feature    82982. .101845
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
     misc_feature    101946. .122755
                     /note="assembly_fragment"
     misc_feature    122856. .142457
                     /note="assembly_fragment"
     misc_feature    142558. .167357
```

```
                     /note="assembly_fragment"
ORIGIN

Alignment Scores:
Pred. No.:                5.31e-62        Length:          167357
Score:                    77.00           Matches:         77
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              5.39%           Indels:          0
DB:                       2               Gaps:            0

US-09-787-097-12 (1-1429) x AC055771 (1-167357)

Qy         985 GluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrp 1004
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      113358 GAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGG 113299

Qy        1005 CysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyPro 1024
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      113298 TGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCA 113239

Qy        1025 ValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSer 1044
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      113238 GTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCC 113179

Qy        1045 SerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysPro 1061
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      113178 AGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCA 113128

RESULT 35
AP002898
LOCUS       AP002898               206647 bp    DNA     linear   PRI 24-OCT-2000
DEFINITION  Homo sapiens genomic DNA, chromosome 20p, clone:13N6, complete
            sequence.
ACCESSION   AP002898
VERSION     AP002898.1  GI:10998163
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 206647)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Homo sapiens 206,647 genomic DNA of 20p
  JOURNAL   Published Only in DataBase (2000)
REFERENCE   2  (bases 1 to 206647)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-OCT-2000) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Overlapping Clones: AL132773, AL109804
            Cytogenetic Position: 20p
            STS Markers: D20S828, D20S813, stSG408, A005O05, sts-H22126,
            D20S867, D20S752
            Additional author information
            Hayflick, S.J.
            Oregon Health Sciences University, Molecular & Medical Genetics
            Dept. 3181 SW Sam Jackson Park Road, MC-L103, BSA-4596, Portland OR
            97201-3098, USA
            phone:1-503-494-6866, fax:1-503-494-6886
            e-mail: hayflick@ohsu.edu.
FEATURES              Location/Qualifiers
     source           1. .206647
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /chromosome="20"
                      /map="20p"
                      /clone="13N6"
ORIGIN

Alignment Scores:
Pred. No.:                6.29e-62        Length:          206647
Score:                    77.00           Matches:         77
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              5.39%           Indels:          0
DB:                       9               Gaps:            0

US-09-787-097-12 (1-1429) x AP002898 (1-206647)

Qy         985 GluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrp 1004
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61935 GAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGG 61994

Qy        1005 CysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyPro 1024
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61995 TGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCA 62054

Qy        1025 ValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSer 1044
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       62055 GTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCC 62114

Qy        1045 SerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysPro 1061
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       62115 AGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCA 62165

RESULT 36
AX326809
LOCUS       AX326809               207433 bp    DNA     linear   PAT 07-JAN-2002
DEFINITION  Sequence 5 from Patent WO0178894.
ACCESSION   AX326809
VERSION     AX326809.1  GI:18097525
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Keith,T.
  TITLE     Novel human gene relating to respiratory diseases, obesity, and
            inflammatory bowel disease
  JOURNAL   Patent: WO 0178894-A 5 25-OCT-2001;
            Genome Therapeutics Corp. (US)
FEATURES              Location/Qualifiers
     source           1. .207433
                      /organism="Homo sapiens"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                6.31e-62        Length:          207433
Score:                    77.00           Matches:         77
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              5.39%           Indels:          0
DB:                       6               Gaps:            0

US-09-787-097-12 (1-1429) x AX326809 (1-207433)

Qy         985 GluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrp 1004
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        9815 GAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGG 9874

Qy        1005 CysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyPro 1024
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        9875 TGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCA 9934

Qy        1025 ValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSer 1044
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       9935 GTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCC 9994

Qy       1045 SerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysPro 1061
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9995 AGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCA 10045

RESULT 37
AR164814
LOCUS       AR164814               1051 bp    DNA     linear   PAT 17-OCT-2001
DEFINITION  Sequence 10 from patent US 6274339.
ACCESSION   AR164814
VERSION     AR164814.1  GI:16238023
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1051)
  AUTHORS   Moore,K. and Nagle,D.Lynn.
  TITLE     Methods and compositions for the diagnosis and treatment of body
            weight disorders, including obesity
  JOURNAL   Patent: US 6274339-A 10 14-AUG-2001;
FEATURES             Location/Qualifiers
     source          1. .1051
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:              8.39e-63        Length:         1051
Score:                  76.00           Matches:        111
Percent Similarity:     98.23%          Conservative:   0
Best Local Similarity:  98.23%          Mismatches:     1
Query Match:            5.32%           Indels:         2
DB:                     6               Gaps:           0

US-09-787-097-12 (1-1429) x AR164814 (1-1051)

Qy         87 AlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGluCysAspArgProCysVal 106
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        463 GCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAAGGAATGTGACCGGCCGTGTGTC 522

Qy        107 AsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCys-ProAlaGlyTrpValGl 126
              ||||||||||||||||||||||||||||||||||||||||| ||| ||||||||||||
Db        523 AACGGCGGCCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCACG-GGCTGGGTGGG 581

Qy        126 yGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPheValTh 146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        582 CGAGCAATGCCAGCACTGCGGGGGGCCGCTTCAGACTAACTGGCTCTTCTGGATTTGTAAC 641

Qy        146 rAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeuIleGluGlyGlnPr 166
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        642 AGATGGACCTGGGAATTATAAATATAAGACGAAGTGCACATGGCTCATTGAAGGACAGCC 701

Qy        166 oAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLe 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        702 AAATAGAATAATGAGACTTCGCTTCAACCATTTTGCTACAGAATGTAGCTGGGACCATTT 761

Qy        186 uTyrValTyrAspGlyAspSerIleTyrAlaProLeu 198
              |||||||||||||||||||||||||||||||||||
Db        762 ATATGTTTATGATGGGGACTCAATCTACGCACCTCTG 798

RESULT 38
H009384S11
LOCUS       H009384S11             3589 bp    DNA     linear   PRI 14-DEC-2000
DEFINITION  Homo sapiens attractin precursor (ATRN) gene, exon 12.
ACCESSION   AF218899
VERSION     AF218899.1  GI:8118065
KEYWORDS    .
SEGMENT     11 of 27
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3589)
  AUTHORS   Tang,W., Gunn,T.M., McLaughlin,D.F., Barsh,G.S., Schlossman,S.F.
            and Duke-Cohan,J.S.
  TITLE     Secreted and membrane attractin result from alternative splicing of
            the human ATRN gene
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
  MEDLINE   20283944
   PUBMED   10811918
REFERENCE   2  (bases 1 to 3589)
  AUTHORS   Tang,W., Gunn,T.M., McLaughlin,D.F., Freeman,G.J., Barsh,G.S.,
            Schlossman,S.F. and Duke-Cohan,J.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
            Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
            USA
FEATURES             Location/Qualifiers
     source          1. .3589
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /map="20p13"
     exon            1585. .1805
                     /gene="ATRN"
                     /number=12
ORIGIN

Alignment Scores:
Pred. No.:              1.89e-60        Length:         3589
Score:                  74.00           Matches:        74
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            5.18%           Indels:         0
DB:                     9               Gaps:           0

US-09-787-097-12 (1-1429) x H009384S11 (1-3589)

Qy        624 SerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPhe 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1583 AGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTC 1642

Qy        644 ThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGly 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1643 ACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGT 1702

Qy        664 IleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAsp 683
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1703 ATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGAT 1762

Qy        684 GluGlnGluGluLysLeuLysSerGluCysPheSerLysArg 697
              |||||||||||||||||||||||||||||||||||||||||||
Db       1763 GAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGA 1804

RESULT 39
AC023420/c
LOCUS       AC023420               155869 bp    DNA     linear   HTG 23-SEP-2000
DEFINITION  Homo sapiens chromosome 2 clone RP11-737G10 map 2, WORKING DRAFT
            SEQUENCE, 15 unordered pieces.
ACCESSION   AC023420
VERSION     AC023420.4  GI:10280876
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 155869)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 2, clone RP11-737G10
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 155869)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beda,F., Boguslavkiy,L.,
            Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
```

```
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Dodge,S., Domino,M., Doyle,M.,
            Fenestor,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., Landers,T., Largocque,K., Lehoczky,J., Levine,R.,
            Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M.,
            McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J.,
            Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Naylor,J.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Olivar,T.M.,
            Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C.,
            Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S.,
            Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,
            Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B.,
            Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and
            Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-FEB-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Sep 23, 2000 this sequence version replaced gi:7767820.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L6805
                Center clone name: 737_G_10
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 145736 bases at least Q40
                Consensus quality: 151534 bases at least Q30
                Consensus quality: 153551 bases at least Q20
                Insert size: 157000; agarose-fp
                Insert size: 154469; sum-of-contigs
                Quality coverage: 4.2 in Q20 bases; agarose-fp
                Quality coverage: 4.3 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 15 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    12284: contig of 12284 bp in length
            *    12285    12384: gap of 100 bp
            *    12385    13704: contig of 1320 bp in length
            *    13705    13804: gap of 100 bp
            *    13805    15657: contig of 1853 bp in length
            *    15658    15757: gap of 100 bp
            *    15758    17755: contig of 1998 bp in length
            *    17756    17855: gap of 100 bp
            *    17856    19875: contig of 2020 bp in length
            *    19876    19975: gap of 100 bp
            *    19976    22894: contig of 2919 bp in length
            *    22895    22994: gap of 100 bp
            *    22995    26992: contig of 3998 bp in length
            *    26993    27092: gap of 100 bp
            *    27093    31422: contig of 4330 bp in length
            *    31423    31522: gap of 100 bp
            *    31523    37920: contig of 6398 bp in length
            *    37921    38020: gap of 100 bp
            *    38021    48644: contig of 10624 bp in length
            *    48645    48744: gap of 100 bp
            *    48745    69112: contig of 20368 bp in length
            *    69113    69212: gap of 100 bp
            *    69213    86049: contig of 16837 bp in length
            *    86050    86149: gap of 100 bp
            *    86150   107477: contig of 21328 bp in length
            *   107478   107577: gap of 100 bp
            *   107578   131094: contig of 23517 bp in length
            *   131095   131194: gap of 100 bp
            *   131195   155869: contig of 24675 bp in length.
FEATURES             Location/Qualifiers
     source          1. .155869
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /map="2"
                     /clone="RP11-737G10"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .12284
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    12385. .13704
                     /note="assembly_fragment"
     misc_feature    13805. .15657
                     /note="assembly_fragment"
     misc_feature    15758. .17755
                     /note="assembly_fragment"
     misc_feature    17856. .19875
                     /note="assembly_fragment"
     misc_feature    19976. .22894
                     /note="assembly_fragment"
     misc_feature    22995. .26992
                     /note="assembly_fragment"
     misc_feature    27093. .31422
                     /note="assembly_fragment"
     misc_feature    31523. .37920
                     /note="assembly_fragment"
     misc_feature    38021. .48644
                     /note="assembly_fragment"
     misc_feature    48745. .69112
                     /note="assembly_fragment"
     misc_feature    69213. .86049
                     /note="assembly_fragment"
     misc_feature    86150. .107477
                     /note="assembly_fragment"
     misc_feature    107578. .131094
                     /note="assembly_fragment"
     misc_feature    131195. .155869
                     /note="assembly_fragment"
ORIGIN

Alignment Scores:
Pred. No.:              3.88e-59        Length:         155869
Score:                  74.00           Matches:        74
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            5.18%           Indels:         0
DB:                     2               Gaps:           0

US-09-787-097-12 (1-1429) x AC023420 (1-155869)

Qy        624 SerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPhe 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10396 AGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTC 10337

Qy        644 ThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGly 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10336 ACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGT 10277

Qy        664 IleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAsp 683
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10276 ATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGAT 10217
```

```
Qy        684 GluGlnGluGluLysLeuLysSerGluCysPheSerLysArg 697
              ||||||||||||||||||||||||||||||||||||||||||||
Db      10216 GAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGA 10175

RESULT 40
AB06968S15
LOCUS       AB06968S15              1207 bp    DNA     linear   ROD 04-APR-2002
DEFINITION  Rattus norvegicus Atrn gene for membrane attractin, exon 16.
ACCESSION   AB049236
VERSION     AB049236.1  GI:12275377
KEYWORDS    .
SEGMENT     15 of 27
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1
  AUTHORS   Kuramoto,T., Kitada,K., Inui,T., Sasaki,Y., Ito,K., Hase,T.,
            Kawaguchi,S., Ogawa,Y., Nakao,K., Barsh,G.S., Nagao,M., Ushijima,T.
            and Serikawa,T.
  TITLE     Attractin/mahogany/zitter plays a critical role in myelination of
            the central nervous system
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (2), 559-564 (2001)
  MEDLINE   21143347
   PUBMED   11209055
REFERENCE   2  (bases 1 to 1207)
  AUTHORS   Kuramoto,T. and Ushijima,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-SEP-2000) Takashi Kuramoto, National Cancer Center
            Research Institute, Carcinogenesis Division; Tsukiji 5-1-1,
            Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramot@gan2.ncc.go.jp,
            Tel:81-33542-2511, Fax:81-5565-1753)
FEATURES             Location/Qualifiers
     source          1. .1207
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /chromosome="3"
     exon            488. .707
                     /gene="Atrn"
                     /product="membrane attractin"
                     /note="CDS is reported in Acc# AB049248"
                     /number=16
ORIGIN

Alignment Scores:
Pred. No.:              7.24e-60        Length:         1207
Score:                  73.00           Matches:        73
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            5.11%           Indels:         0
DB:                     10              Gaps:           0

US-09-787-097-12 (1-1429) x AB06968S15 (1-1207)

Qy        849 SerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 TCCAAGCTCACTCTGACTCCATGGGTTGGTCTTCGGAAGATCAATGTATCTTACTGGTGC 547

Qy        869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 TGGGAGGATATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGT 607

Qy        889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        608 GATGCCGGCTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGACTGAAGGCTGCAACC 667

Qy        909 CysIleAsnProLeuAsnGlySerValCysGluArgPro 921
              |||||||||||||||||||||||||||||||||||||
Db        668 TGCATCAACCCACTCAATGGAAGTGTTTGTGAAAGGCCT 706

RESULT 41
H009384S14
LOCUS       H009384S14              1728 bp    DNA     linear   PRI 14-DEC-2000
DEFINITION  Homo sapiens attractin precursor (ATRN) gene, exon 16.
ACCESSION   AF218902
VERSION     AF218902.1  GI:8118068
KEYWORDS    .
SEGMENT     14 of 27
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1728)
  AUTHORS   Tang,W., Gunn,T.M., McLaughlin,D.F., Barsh,G.S., Schlossman,S.F.
            and Duke-Cohan,J.S.
  TITLE     Secreted and membrane attractin result from alternative splicing of
            the human ATRN gene
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
  MEDLINE   20283944
   PUBMED   10811918
REFERENCE   2  (bases 1 to 1728)
  AUTHORS   Tang,W., Gunn,T.M., McLaughlin,D.F., Freeman,G.J., Barsh,G.S.,
            Schlossman,S.F. and Duke-Cohan,J.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
            Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
            USA
FEATURES             Location/Qualifiers
     source          1. .1728
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /map="20p13"
     exon            533. .752
                     /gene="ATRN"
                     /number=16
ORIGIN

Alignment Scores:
Pred. No.:              9.65e-60        Length:         1728
Score:                  73.00           Matches:        73
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            5.11%           Indels:         0
DB:                     9               Gaps:           0

US-09-787-097-12 (1-1429) x H009384S14 (1-1728)

Qy        849 SerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 TCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGC 592

Qy        869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 TGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGT 652

Qy        889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 GATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACC 712

Qy        909 CysIleAsnProLeuAsnGlySerValCysGluArgPro 921
              |||||||||||||||||||||||||||||||||||||
Db        713 TGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCT 751

RESULT 42
AR164810
LOCUS       AR164810                90050 bp    DNA     linear   PAT 17-OCT-2001
DEFINITION  Sequence 5 from patent US 6274339.
ACCESSION   AR164810
VERSION     AR164810.1  GI:16238016
KEYWORDS    .
SOURCE      Unknown.
```

```
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 90050)
  AUTHORS   Moore,K. and Nagle,D.Lynn.
  TITLE     Methods and compositions for the diagnosis and treatment of body
            weight disorders, including obesity
  JOURNAL   Patent: US 6274339-A 5 14-AUG-2001;
FEATURES             Location/Qualifiers
     source          1. .90050
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:                2.29e-58        Length:         90050
Score:                    73.00           Matches:        73
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              5.11%           Indels:         0
DB:                       6               Gaps:           0

US-09-787-097-12 (1-1429) x AR164810 (1-90050)

Qy        849 SerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      28996 TCCAAGCTCACTCTGACTCCATGGGTTGGTCTTCGGAAGATCAATGTGTCTTACTGGTGC 29055

Qy        869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      29056 TGGGAGGATATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGT 29115

Qy        889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      29116 GATGCTGGCTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGATTAAAGGCTGCAACC 29175

Qy        909 CysIleAsnProLeuAsnGlySerValCysGluArgPro 921
              |||||||||||||||||||||||||||||||||||||||
Db      29176 TGCATCAACCCTCTCAATGGCAGCGTCTGTGAAAGGCCT 29214

RESULT 43
AL833771
LOCUS       AL833771              189283 bp    DNA     linear   ROD 20-AUG-2002
DEFINITION  Mouse DNA sequence from clone RP23-228H14 on chromosome 2, complete
            sequence.
ACCESSION   AL833771
VERSION     AL833771.4  GI:22449827
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 189283)
  AUTHORS   Williams,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-AUG-2002) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Aug 22, 2002 this sequence version replaced gi:22416184.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-228H14 is
            from the RPCI-23 Mouse PAC Library
            constructed by the group of Pieter de Jong.
            For further details see http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6.
FEATURES             Location/Qualifiers
     source          1. .189283
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="2"
                     /clone="RP23-228H14"
                     /clone_lib="RPCI-23"
ORIGIN

Alignment Scores:
Pred. No.:                4.16e-58        Length:         189283
Score:                    73.00           Matches:        73
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              5.11%           Indels:         0
DB:                       10              Gaps:           0

US-09-787-097-12 (1-1429) x AL833771 (1-189283)

Qy        849 SerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      98835 TCCAAGCTCACTCTGACTCCATGGGTTGGTCTTCGGAAGATCAATGTGTCTTACTGGTGC 98894

Qy        869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      98895 TGGGAGGATATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGT 98954

Qy        889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      98955 GATGCTGGCTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGATTAAAGGCTGCAACC 99014

Qy        909 CysIleAsnProLeuAsnGlySerValCysGluArgPro 921
              |||||||||||||||||||||||||||||||||||||||
Db      99015 TGCATCAACCCTCTCAATGGCAGCGTCTGTGAAAGGCCT 99053

RESULT 44
AC094932/c
LOCUS       AC094932              245589 bp    DNA     linear   HTG 09-MAY-2003
DEFINITION  Rattus norvegicus clone CH230-6D19, *** SEQUENCING IN PROGRESS ***.
ACCESSION   AC094932
VERSION     AC094932.5  GI:30466943
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 245589)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
```

Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 245589)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (15-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 245589)
AUTHORS Rat Genome Sequencing Consortium.
TITLE Direct Submission
JOURNAL Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On May 9, 2003 this sequence version replaced gi:23264765.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.
---------- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
---------- Project Information
Center project name: GBTI
Center clone name: CH230-6D19

---------- Summary Statistics
Assembly program: Atlas;
Consensus quality: 205514 bases at least Q40
Consensus quality: 207203 bases at least Q30
Consensus quality: 208050 bases at least Q20
Estimated insert size: 221727; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 245589: contig of 245589 bp in length.

FEATURES Location/Qualifiers
source 1. .245589
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-6D19"
misc_feature 545. .1006
/note="clone_boundary
clone_end:Sp6
site:EcoRI
end_sequence:BH358667"
misc_feature 209973. .210849
/note="clone_boundary
clone_end:T7
site:EcoRI
end_sequence:BH358665"
ORIGIN

Alignment Scores:

Pred. No.:	5.12e-58	Length:	245589
Score:	73.00	Matches:	73
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.11%	Indels:	0
DB:	2	Gaps:	0

US-09-787-097-12 (1-1429) x AC094932 (1-245589)

```
Qy        849 SerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      99645 TCCAAGCTCACTCTGACTCCATGGGTTGGTCTTCGGAAGATCAATGTATCTTACTGGTGC 99586

Qy        869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      99585 TGGGAGGATATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGT 99526

Qy        889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      99525 GATGCCGGCTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGACTGAAGGCTGCAACC 99466

Qy        909 CysIleAsnProLeuAsnGlySerValCysGluArgPro 921
              |||||||||||||||||||||||||||||||||||||||
Db      99465 TGCATCAACCCACTCAATGGAAGTGTTTGTGAAAGGCCT 99427
```

RESULT 45
H009384S08
LOCUS H009384S08 2018 bp DNA linear PRI 14-DEC-2000
DEFINITION Homo sapiens attractin precursor (ATRN) gene, exon 8.
ACCESSION AF218896
VERSION AF218896.1 GI:8118062
KEYWORDS .
SEGMENT 8 of 27
SOURCE Homo sapiens (human)

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2018)
  AUTHORS   Tang,W., Gunn,T.M., McLaughlin,D.F., Barsh,G.S., Schlossman,S.F.
            and Duke-Cohan,J.S.
  TITLE     Secreted and membrane attractin result from alternative splicing of
            the human ATRN gene
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
  MEDLINE   20283944
   PUBMED   10811918
REFERENCE   2  (bases 1 to 2018)
  AUTHORS   Tang,W., Gunn,T.M., McLaughlin,D.F., Freeman,G.J., Barsh,G.S.,
            Schlossman,S.F. and Duke-Cohan,J.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
            Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
            USA
FEATURES             Location/Qualifiers
     source          1. .2018
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /map="20p13"
     exon            926. .1169
                     /gene="ATRN"
                     /number=8
ORIGIN

Alignment Scores:
Pred. No.:              7.76e-56        Length:         2018
Score:                  69.00           Matches:        69
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            4.83%           Indels:         0
DB:                     9               Gaps:           0

US-09-787-097-12 (1-1429) x H009384S08 (1-2018)

Qy        414 ThrGlyAsnValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeu 433
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        962 ACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTG 1021

Qy        434 LeuThrProLysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThr 453
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1022 TTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACA 1081

Qy        454 LeuLysAsnGlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyr 473
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1082 CTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATAT 1141

Qy        474 IleSerAsnValGlnGluTyrAspLeu 482
              |||||||||||||||||||||||||||
Db       1142 ATAAGCAATGTGCAGGAATATGATTTG 1168

RESULT 46
H009384S01
LOCUS       H009384S01              1576 bp    DNA     linear   PRI 14-DEC-2000
DEFINITION  Homo sapiens attractin precursor (ATRN) gene, exon 1.
ACCESSION   AF218889
VERSION     AF218889.1  GI:8118055
KEYWORDS    .
SEGMENT     1 of 27
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1576)
  AUTHORS   Tang,W., Gunn,T.M., McLaughlin,D.F., Barsh,G.S., Schlossman,S.F.
            and Duke-Cohan,J.S.
  TITLE     Secreted and membrane attractin result from alternative splicing of
            the human ATRN gene
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
  MEDLINE   20283944
   PUBMED   10811918
REFERENCE   2  (bases 1 to 1576)
  AUTHORS   Tang,W., Gunn,T.M., McLaughlin,D.F., Freeman,G.J., Barsh,G.S.,
            Schlossman,S.F. and Duke-Cohan,J.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
            Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
            USA
FEATURES             Location/Qualifiers
     source          1. .1576
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /map="20p13"
     exon            745. .1233
                     /gene="ATRN"
                     /number=1
     sig_peptide     824. .1072
                     /gene="ATRN"
ORIGIN

Alignment Scores:
Pred. No.:              5.84e-55        Length:         1576
Score:                  68.00           Matches:        136
Percent Similarity:     98.55%          Conservative:   0
Best Local Similarity:  98.55%          Mismatches:     1
Query Match:            4.76%           Indels:         2
DB:                     9               Gaps:           0

US-09-787-097-12 (1-1429) x H009384S01 (1-1576)

Qy          1 MetValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        824 ATGGTGGCTGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 883

Qy         21 AlaLeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArg 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        884 GCGCTCGCGGGGCAGGAGCGGCGGGCCGCACTGGGACTGGGACGTGACCAGGGCTGGGAGG 943

Qy         41 ProGlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSerProProLeuArgProArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        944 CCGGGGCTGGGGGCCGGGCTGCGCCTCCCGCGGCTGCTGTCTCCACCGCTGCGGCCACGG 1003

Qy         61 LeuLeuLeuLeuLeuLeuLeuPro-ProProLeuLeuLeuLeuLeuProCysGl 80
              |||||||||||||||||||||||    ||||||||||||||||||||||||||||
Db       1004 CTGCTGCTGCTGCTGTTGTTGCTCTC-GCCGCCGCTGCTGCTGCTGCTGCTGCCCTGTGA 1062

Qy         80 uAlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGl 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1063 GGCCGAGGCCGCGGCGGCGGCGGCGGCGGTGTCGGGCTCAGCCGCAGCCGAGGCCAAGGA 1122

Qy        100 uCysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCy 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1123 ATGTGACCGGCCCTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTG 1182

Qy        120 sProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPheArg 137
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1183 CCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGG 1234

RESULT 47
AL353193
LOCUS       AL353193                51371 bp    DNA     linear   PRI 19-JUL-2001
DEFINITION  Human DNA sequence from clone RP11-137F15 on chromosome 20 Contains
            ESTs, STSs, GSSs and CpG islands. Contains a pseudogene similar to
            a TCEAL1 (transcription elongation factor A (SII)-like 1) homolog
            and the 5' part of the ATRN gene for attractin with
            dipeptidylpeptidase IV activity, complete sequence.
ACCESSION   AL353193
VERSION     AL353193.7  GI:10443410
KEYWORDS    HTG; ATRN; attractin; CpG island; dipeptidylpeptidase IV; TCEAL1;
```

transcription elongation.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 51371)
AUTHORS Ellington,A.
TITLE Direct Submission
JOURNAL Submitted (19-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT On Oct 1, 2000 this sequence version replaced gi:9930864.
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr20
IMPORTANT: This sequence is not the entire insert of clone RP11-137F15 It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.
The true left end of clone RP4-581P3 is at 51272 in this sequence.
The true right end of clone RP4-687F11 is at 100 in this sequence.
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP11-137F15 is from the library RPCI-11.1 constructed by the group of Pieter de Jong. For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6.

FEATURES Location/Qualifiers

source 1. .51371
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/clone="RP11-137F15"
/clone_lib="RPCI-11.1"

repeat_region 11. .374
/note="L1PB3 repeat: matches 5787. .6143 of consensus"

repeat_region 483. .562
/note="40 copies 2 mer ta 63% conserved"

repeat_region 489. .660
/note="43 copies 4 mer tata 57% conserved"

repeat_region 620. .661
/note="21 copies 2 mer at 81% conserved"

misc_feature 3375. .3794
/note="match: GSS: Em:AQ755239"

repeat_region 4861. .5206
/note="MER41D repeat: matches 169. .548 of consensus"

repeat_region 5660. .5699
/note="MER41D repeat: matches 101. .140 of consensus"

repeat_region 5943. .6017
/note="MER41D repeat: matches 1. .75 of consensus"

repeat_region 15032. .15083
/note="LTR37B repeat: matches 233. .284 of consensus"

repeat_region 15142. .15189
/note="12 copies 4 mer acac 97% conserved"

repeat_region 15489. .15537
/note="MER41D repeat: matches 215. .263 of consensus"

repeat_region 15842. .16057
/note="MER41D repeat: matches 1. .215 of consensus"

repeat_region 16420. .16463
/note="11 copies 4 mer caca 90% conserved"

repeat_region 16421. .16462
/note="21 copies 2 mer ac 92% conserved"

misc_feature 18515. .19187
/note="CpG island"
/evidence=not_experimental

repeat_region 22182. .22416
/note="L1MEc repeat: matches 1900. .2132 of consensus"

repeat_region 22722. .23003
/note="L1MEc repeat: matches 2132. .2079 of consensus"

repeat_region 23143. .23183
/note="L1MEc repeat: matches 2078. .2117 of consensus"

misc_feature complement(29399. .29721)
/note="match: GSS: Em:AQ197492"

misc_feature 30460. .30862
/note="match: GSS: Em:AQ720519"

repeat_region 31259. .31479
/note="MER30 repeat: matches 1. .230 of consensus"

repeat_region 33494. .33641
/note="37 copies 4 mer aagg 72% conserved"

repeat_region 35991. .36050
/note="15 copies 4 mer tata 73% conserved"

repeat_region 37080. .37131
/note="26 copies 2 mer gt 98% conserved"

repeat_region 39161. .39359
/note="MER63A repeat: matches 1. .203 of consensus"

repeat_region 42452. .42473
/note="11 copies 2 mer tt 100% conserved"

repeat_region 42475. .42682
/note="L1MEc repeat: matches 2151. .2369 of consensus"

gene complement(42686. .42953)
/gene="bA137F15.1"
/pseudo

CDS complement(42686. .42953)
/gene="bA137F15.1"
/note="bA137F15.1 (novel pseudogene similar to a TCEAL1 (transcription elongation factor A (SII)-like 1) homolog)
match: cDNAs: Em:AF125535 Em:U92454 Em:M99701
match: ESTs: Em:BE564937 Em:BE567354 Em:BE564574 Em:BE565300 Em:BE789934 Em:BE569184 Em:AA031368 Em:BE866689 Em:BE865705 Em:BE565476 Em:BE567379 Em:BE565985 Em:W47079 Em:AA096079 Em:R70575 Em:AI692995 Em:AI832505 Em:AI815035
match: proteins: Tr:Q9UHQ7 Tr:Q15170 Tr:Q9UJQ9"
/pseudo
/codon_start=1
/evidence=not_experimental

repeat_region 42971. .43050
/note="40 copies 2 mer tt 66% conserved"

repeat_region 43002. .43053
/note="13 copies 4 mer ttct 73% conserved"

misc_feature 44223. .44763
/note="match: STS: Em:HS652K19T"

misc_feature 48154. .49093
/note="CpG island"
/evidence=not_experimental

repeat_region 48427. .48470
/note="22 copies 2 mer gc 75% conserved"

repeat_region 48430. .48469
/note="10 copies 4 mer cgca 80% conserved"

gene 48549. .49026
/gene="ATRN"

mRNA 48549. .>49026
/gene="ATRN"
/product="bA137F15.2 (attractin (with dipeptidylpeptidase IV activity))"
/note="match: cDNAs: Em:AF106861 Em:AF034957 Em:AF119821 Em:AF116897"
/evidence=not_experimental

CDS 48617. .>49026

```
                     /gene="ATRN"
                     /note="continues in Em:AL109805 as dJ581P3.2
                     match: proteins: Tr:O75882 Tr:O95414 Tr:Q9WU77 Tr:Q9WU60"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="bA137F15.2 (attractin (with dipeptidylpeptidase
                     IV activity))"
                     /protein_id="CAC17563.1"
                     /db_xref="GI:11493257"
                     /db_xref="SPTREMBL:Q9H1P8"
                     /translation="MVAAAAATEARLRRRTAATAALAGRSGGPHWDWDVTRAGRPGLG
                     AGLRLPRLLSPPLRPRLLLLLLLLSPPLLLLLLPCEAEAAAAAAAVSGSAAAEAKECD
                     RPCVNGGRCNPGTGQCVCPAGWVGEQCQHCGGRF"
     repeat_region   48760. .48849
                     /note="30 copies 3 mer gct 75% conserved"
     repeat_region   48862. .48909
                     /note="16 copies 3 mer ggc 77% conserved"
ORIGIN

Alignment Scores:
Pred. No.:              9.52e-54        Length:         51371
Score:                  68.00           Matches:        136
Percent Similarity:     98.55%          Conservative:   0
Best Local Similarity:  98.55%          Mismatches:     1
Query Match:            4.76%           Indels:         2
DB:                     9               Gaps:           0

US-09-787-097-12 (1-1429) x AL353193 (1-51371)

Qy          1 MetValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48617 ATGGTGGCTGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 48676

Qy         21 AlaLeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArg 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48677 GCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGGGACTGGGACGTGACCAGGGCTGGGAGG 48736

Qy         41 ProGlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSerProProLeuArgProArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48737 CCGGGGCTGGGGGCCGGGCTGCGCCTCCCGCGGCTGCTGTCTCCACCGCTGCGGCCACGG 48796

Qy         61 LeuLeuLeuLeuLeuLeuLeuPro-ProProLeuLeuLeuLeuLeuProCysGl 80
              |||||||||||||||||||||||    ||||||||||||||||||||||||||||
Db      48797 CTGCTGCTGCTGCTGTTGTTGCTCTC-GCCGCCGCTGCTGCTGCTGCTGCCCTGTGA 48855

Qy         80 uAlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGl 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48856 GGCCGAGGCCGCGGCGGCGGCGGCGGCGGTGTCGGGCTCAGCCGCAGCCGAGGCCAAGGA 48915

Qy        100 uCysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCy 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48916 ATGTGACCGGCCCTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTG 48975

Qy        120 sProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPheArg 137
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48976 CCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGG 49027

RESULT 48
AC074061/c
LOCUS       AC074061               63513 bp    DNA     linear   HTG 13-JUL-2000
DEFINITION  Homo sapiens chromosome 2 clone RP11-770E8 map 2, LOW-PASS SEQUENCE
            SAMPLING.
ACCESSION   AC074061
VERSION     AC074061.1  GI:9150468
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 63513)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 2, clone RP11-770E8
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 63513)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JUL-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L5682
                Center clone name: 770_E_8
            --------------.
            * NOTE: This record contains 72 individual
            * sequencing reads that have not been assembled into
            * contigs. Runs of N are used to separate the reads
            * and the order in which they appear is completely
            * arbitrary. Low-pass sequence sampling is useful for
            * identifying clones that may be gene-rich and allows
            * overlap relationships among clones to be deduced.
            * However, it should not be assumed that this clone
            * will be sequenced to completion. In the event that
            * the record is updated, the accession number will
            * be preserved.
            *        1      780: contig of 780 bp in length
            *      781      880: gap of 100 bp
            *      881     1658: contig of 778 bp in length
            *     1659     1758: gap of 100 bp
            *     1759     2556: contig of 798 bp in length
            *     2557     2656: gap of 100 bp
            *     2657     3437: contig of 781 bp in length
            *     3438     3537: gap of 100 bp
            *     3538     4308: contig of 771 bp in length
            *     4309     4408: gap of 100 bp
            *     4409     5165: contig of 757 bp in length
            *     5166     5265: gap of 100 bp
            *     5266     6031: contig of 766 bp in length
            *     6032     6131: gap of 100 bp
            *     6132     6896: contig of 765 bp in length
            *     6897     6996: gap of 100 bp
            *     6997     7771: contig of 775 bp in length
            *     7772     7871: gap of 100 bp
            *     7872     8651: contig of 780 bp in length
            *     8652     8751: gap of 100 bp
            *     8752     9524: contig of 773 bp in length
            *     9525     9624: gap of 100 bp
            *     9625    10447: contig of 823 bp in length
            *    10448    10547: gap of 100 bp
            *    10548    11406: contig of 859 bp in length
```

```
*    11407    11506: gap of 100 bp
*    11507    12265: contig of 759 bp in length
*    12266    12365: gap of 100 bp
*    12366    13122: contig of 757 bp in length
*    13123    13222: gap of 100 bp
*    13223    13980: contig of 758 bp in length
*    13981    14080: gap of 100 bp
*    14081    14922: contig of 842 bp in length
*    14923    15022: gap of 100 bp
*    15023    15805: contig of 783 bp in length
*    15806    15905: gap of 100 bp
*    15906    16664: contig of 759 bp in length
*    16665    16764: gap of 100 bp
*    16765    17537: contig of 773 bp in length
*    17538    17637: gap of 100 bp
*    17638    18409: contig of 772 bp in length
*    18410    18509: gap of 100 bp
*    18510    19278: contig of 769 bp in length
*    19279    19378: gap of 100 bp
*    19379    20137: contig of 759 bp in length
*    20138    20237: gap of 100 bp
*    20238    21012: contig of 775 bp in length
*    21013    21112: gap of 100 bp
*    21113    21892: contig of 780 bp in length
*    21893    21992: gap of 100 bp
*    21993    22810: contig of 818 bp in length
*    22811    22910: gap of 100 bp
*    22911    23712: contig of 802 bp in length
*    23713    23812: gap of 100 bp
*    23813    24594: contig of 782 bp in length
*    24595    24694: gap of 100 bp
*    24695    25452: contig of 758 bp in length
*    25453    25552: gap of 100 bp
*    25553    26318: contig of 766 bp in length
*    26319    26418: gap of 100 bp
*    26419    27200: contig of 782 bp in length
*    27201    27300: gap of 100 bp
*    27301    28095: contig of 795 bp in length
*    28096    28195: gap of 100 bp
*    28196    28977: contig of 782 bp in length
*    28978    29077: gap of 100 bp
*    29078    29883: contig of 806 bp in length
*    29884    29983: gap of 100 bp
*    29984    30776: contig of 793 bp in length
*    30777    30876: gap of 100 bp
*    30877    31661: contig of 785 bp in length
*    31662    31761: gap of 100 bp
*    31762    32517: contig of 756 bp in length
*    32518    32617: gap of 100 bp
*    32618    33392: contig of 775 bp in length
*    33393    33492: gap of 100 bp
*    33493    34254: contig of 762 bp in length
*    34255    34354: gap of 100 bp
*    34355    35148: contig of 794 bp in length
*    35149    35248: gap of 100 bp
*    35249    36040: contig of 792 bp in length
*    36041    36140: gap of 100 bp
*    36141    36912: contig of 772 bp in length
*    36913    37012: gap of 100 bp
*    37013    37807: contig of 795 bp in length
*    37808    37907: gap of 100 bp
*    37908    38664: contig of 757 bp in length
*    38665    38764: gap of 100 bp
*    38765    39602: contig of 838 bp in length
*    39603    39702: gap of 100 bp
*    39703    40477: contig of 775 bp in length
*    40478    40577: gap of 100 bp
*    40578    41343: contig of 766 bp in length
*    41344    41443: gap of 100 bp
*    41444    42203: contig of 760 bp in length
*    42204    42303: gap of 100 bp
*    42304    43067: contig of 764 bp in length
*    43068    43167: gap of 100 bp
*    43168    43963: contig of 796 bp in length
*    43964    44063: gap of 100 bp
*    44064    44885: contig of 822 bp in length
*    44886    44985: gap of 100 bp
*    44986    45745: contig of 760 bp in length
*    45746    45845: gap of 100 bp
*    45846    46615: contig of 770 bp in length
*    46616    46715: gap of 100 bp
*    46716    47482: contig of 767 bp in length
*    47483    47582: gap of 100 bp
*    47583    48371: contig of 789 bp in length
*    48372    48471: gap of 100 bp
*    48472    49248: contig of 777 bp in length
*    49249    49348: gap of 100 bp
*    49349    50123: contig of 775 bp in length
*    50124    50223: gap of 100 bp
*    50224    51000: contig of 777 bp in length
*    51001    51100: gap of 100 bp
*    51101    51857: contig of 757 bp in length
*    51858    51957: gap of 100 bp
*    51958    52818: contig of 861 bp in length
*    52819    52918: gap of 100 bp
*    52919    53694: contig of 776 bp in length
*    53695    53794: gap of 100 bp
*    53795    54564: contig of 770 bp in length
*    54565    54664: gap of 100 bp
*    54665    55453: contig of 789 bp in length
*    55454    55553: gap of 100 bp
*    55554    56424: contig of 871 bp in length
*    56425    56524: gap of 100 bp
*    56525    57278: contig of 754 bp in length
*    57279    57378: gap of 100 bp
*    57379    58156: contig of 778 bp in length
*    58157    58256: gap of 100 bp
*    58257    59009: contig of 753 bp in length
*    59010    59109: gap of 100 bp
*    59110    59903: contig of 794 bp in length
*    59904    60003: gap of 100 bp
*    60004    60783: contig of 780 bp in length
*    60784    60883: gap of 100 bp

Alignment Scores:
Pred. No.:                1.13e-53        Length:          63513
Score:                    68.00           Matches:         68
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              4.76%           Indels:          0
DB:                       2               Gaps:            0

US-09-787-097-12 (1-1429) x AC074061 (1-63513)

Qy         70 ProProLeuLeuLeuLeuLeuProCysGluAlaGluAlaAlaAlaAlaAlaAlaAla 89
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      31457 CCGCCGCTGCTGCTGCTGCTGCCCTGTGAGGCCGAGGCCGCGGCGGCGGCGGCGCG 31398

Qy         90 ValSerGlySerAlaAlaAlaGluAlaLysGluCysAspArgProCysValAsnGlyGly 109
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      31397 GTGTCGGGCTCAGCCGCAGCCGAGGCCAAGGAATGTGACCGGCCCTGTGTCAACGGCGGT 31338

Qy        110 ArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTrpValGlyGluGlnCys 129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      31337 CGCTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGCGAGCAATGC 31278

Qy        130 GlnHisCysGlyGlyArgPheArg 137
              |||||||||||||||||||||||
Db      31277 CAGCACTGCGGGGGCCGCTTCAGG 31254

RESULT 49
AC023420
LOCUS       AC023420               155869 bp    DNA     linear   HTG 23-SEP-2000
DEFINITION  Homo sapiens chromosome 2 clone RP11-737G10 map 2, WORKING DRAFT
            SEQUENCE, 15 unordered pieces.
ACCESSION   AC023420
```

```
VERSION     AC023420.4  GI:10280876
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 155869)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 2, clone RP11-737G10
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 155869)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beda,F., Boguslavkiy,L.,
            Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Dodge,S., Domino,M., Doyle,M.,
            Fenestor,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., Landers,T., Largocque,K., Lehoczky,J., Levine,R.,
            Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M.,
            McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J.,
            Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Naylor,J.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Olivar,T.M.,
            Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C.,
            Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S.,
            Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,
            Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B.,
            Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and
            Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-FEB-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Sep 23, 2000 this sequence version replaced gi:7767820.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L6805
                Center clone name: 737_G_10
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 145736 bases at least Q40
                Consensus quality: 151534 bases at least Q30
                Consensus quality: 153551 bases at least Q20
                Insert size: 157000; agarose-fp
                Insert size: 154469; sum-of-contigs
                Quality coverage: 4.2 in Q20 bases; agarose-fp
                Quality coverage: 4.3 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 15 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    12284: contig of 12284 bp in length
            *    12285    12384: gap of 100 bp
            *    12385    13704: contig of 1320 bp in length
            *    13705    13804: gap of 100 bp
            *    13805    15657: contig of 1853 bp in length
            *    15658    15757: gap of 100 bp
            *    15758    17755: contig of 1998 bp in length
            *    17756    17855: gap of 100 bp
            *    17856    19875: contig of 2020 bp in length
            *    19876    19975: gap of 100 bp
            *    19976    22894: contig of 2919 bp in length
            *    22895    22994: gap of 100 bp
            *    22995    26992: contig of 3998 bp in length
            *    26993    27092: gap of 100 bp
            *    27093    31422: contig of 4330 bp in length
            *    31423    31522: gap of 100 bp
            *    31523    37920: contig of 6398 bp in length
            *    37921    38020: gap of 100 bp
            *    38021    48644: contig of 10624 bp in length
            *    48645    48744: gap of 100 bp
            *    48745    69112: contig of 20368 bp in length
            *    69113    69212: gap of 100 bp
            *    69213    86049: contig of 16837 bp in length
            *    86050    86149: gap of 100 bp
            *    86150   107477: contig of 21328 bp in length
            *   107478   107577: gap of 100 bp
            *   107578   131094: contig of 23517 bp in length
            *   131095   131194: gap of 100 bp
            *   131195   155869: contig of 24675 bp in length.
FEATURES             Location/Qualifiers
     source          1. .155869
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /map="2"
                     /clone="RP11-737G10"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .12284
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    12385. .13704
                     /note="assembly_fragment"
     misc_feature    13805. .15657
                     /note="assembly_fragment"
     misc_feature    15758. .17755
                     /note="assembly_fragment"
     misc_feature    17856. .19875
                     /note="assembly_fragment"
     misc_feature    19976. .22894
                     /note="assembly_fragment"
     misc_feature    22995. .26992
                     /note="assembly_fragment"
     misc_feature    27093. .31422
                     /note="assembly_fragment"
     misc_feature    31523. .37920
                     /note="assembly_fragment"
     misc_feature    38021. .48644
                     /note="assembly_fragment"
     misc_feature    48745. .69112
                     /note="assembly_fragment"
     misc_feature    69213. .86049
                     /note="assembly_fragment"
     misc_feature    86150. .107477
                     /note="assembly_fragment"
     misc_feature    107578. .131094
                     /note="assembly_fragment"
     misc_feature    131195. .155869
                     /note="assembly_fragment"
ORIGIN

Alignment Scores:
Pred. No.:              2.32e-53        Length:         155869
Score:                  68.00           Matches:        136
Percent Similarity:     98.55%          Conservative:   0
Best Local Similarity:  98.55%          Mismatches:     1
Query Match:            4.76%           Indels:         2
DB:                     2               Gaps:           0
```

```
US-09-787-097-12 (1-1429) x AC023420 (1-155869)

Qy          1 MetValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142642 ATGGTGGCTGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 142701

Qy         21 AlaLeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArg 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142702 GCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGGGACTGGGACGTGACCAGGGCTGGGAGG 142761

Qy         41 ProGlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSerProProLeuArgProArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142762 CCGGGGCTGGGGGCCGGGCTGCGCCTCCCGCGGCTGCTGTCTCCACCGCTGCGGCCACGG 142821

Qy         61 LeuLeuLeuLeuLeuLeuLeuPro-ProProLeuLeuLeuLeuLeuProCysGl 80
              ||||||||||||||||||||||    |||||||||||||||||||||||||||||
Db     142822 CTGCTGCTGCTGCTGTTGTTGCTCTC-GCCGCCGCTGCTGCTGCTGCTGCCCTGTGA 142880

Qy         80 uAlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGl 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142881 GGCCGAGGCCGCGGCGGCGGCGGCGGCGGTGTCGGGCTCAGCCGCAGCCGAGGCCAAGGA 142940

Qy        100 uCysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCy 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142941 ATGTGACCGGCCCTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTG 143000

Qy        120 sProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPheArg 137
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143001 CCCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGG 143052

RESULT 50
H009384S13
LOCUS       H009384S13              3112 bp    DNA     linear   PRI 14-DEC-2000
DEFINITION  Homo sapiens attractin precursor (ATRN) gene, exons 14 and 15.
ACCESSION   AF218901
VERSION     AF218901.1  GI:8118067
KEYWORDS    .
SEGMENT     13 of 27
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3112)
  AUTHORS   Tang,W., Gunn,T.M., McLaughlin,D.F., Barsh,G.S., Schlossman,S.F.
            and Duke-Cohan,J.S.
  TITLE     Secreted and membrane attractin result from alternative splicing of
            the human ATRN gene
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
  MEDLINE   20283944
   PUBMED   10811918
REFERENCE   2  (bases 1 to 3112)
  AUTHORS   Tang,W., Gunn,T.M., McLaughlin,D.F., Freeman,G.J., Barsh,G.S.,
            Schlossman,S.F. and Duke-Cohan,J.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
            Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
            USA
FEATURES             Location/Qualifiers
     source          1. .3112
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /map="20p13"
     exon            440. .578
                     /gene="ATRN"
                     /number=14
     exon            2163. .2353
                     /gene="ATRN"
                     /number=15
ORIGIN
```

```
Alignment Scores:
Pred. No.:                7.14e-51      Length:          3112
Score:                    64.00         Matches:         64
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              4.48%         Indels:          0
DB:                       9             Gaps:            0

US-09-787-097-12 (1-1429) x H009384S13 (1-3112)

Qy        785 GluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThrThr 804
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2162 GAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACT 2221

Qy        805 AlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAla 824
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2222 GCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCT 2281

Qy        825 SerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGlnSer 844
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2282 TCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCA 2341

Qy        845 SerGlnSerMet 848
              ||||||||||||
Db       2342 TCTCAGAGCATG 2353

RESULT 51
AC013324
LOCUS       AC013324              118540 bp    DNA     linear   HTG 13-JUL-2000
DEFINITION  Homo sapiens chromosome 2 clone RP11-388K24 map 2, LOW-PASS
            SEQUENCE SAMPLING.
ACCESSION   AC013324
VERSION     AC013324.3  GI:9123914
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 118540)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 2, clone RP11-388K24
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 118540)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baldwin,J., Barna,N., Beckerly,R., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
            Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
            Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
            McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
            Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
            Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-NOV-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Jul 13, 2000 this sequence version replaced gi:6289165.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L3787
                Center clone name: 388_K_24
            --------------.
```

```
* NOTE: This record contains 125 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
*        1      856: contig of 856 bp in length
*      857      956: gap of 100 bp
*      957     1821: contig of 865 bp in length
*     1822     1921: gap of 100 bp
*     1922     2776: contig of 855 bp in length
*     2777     2876: gap of 100 bp
*     2877     3767: contig of 891 bp in length
*     3768     3867: gap of 100 bp
*     3868     4709: contig of 842 bp in length
*     4710     4809: gap of 100 bp
*     4810     5655: contig of 846 bp in length
*     5656     5755: gap of 100 bp
*     5756     6627: contig of 872 bp in length
*     6628     6727: gap of 100 bp
*     6728     7586: contig of 859 bp in length
*     7587     7686: gap of 100 bp
*     7687     8541: contig of 855 bp in length
*     8542     8641: gap of 100 bp
*     8642     9524: contig of 883 bp in length
*     9525     9624: gap of 100 bp
*     9625    10509: contig of 885 bp in length
*    10510    10609: gap of 100 bp
*    10610    11474: contig of 865 bp in length
*    11475    11574: gap of 100 bp
*    11575    12441: contig of 867 bp in length
*    12442    12541: gap of 100 bp
*    12542    13409: contig of 868 bp in length
*    13410    13509: gap of 100 bp
*    13510    14439: contig of 930 bp in length
*    14440    14539: gap of 100 bp
*    14540    15433: contig of 894 bp in length
*    15434    15533: gap of 100 bp
*    15534    16442: contig of 909 bp in length
*    16443    16542: gap of 100 bp
*    16543    17421: contig of 879 bp in length
*    17422    17521: gap of 100 bp
*    17522    18380: contig of 859 bp in length
*    18381    18480: gap of 100 bp
*    18481    19345: contig of 865 bp in length
*    19346    19445: gap of 100 bp
*    19446    20327: contig of 882 bp in length
*    20328    20427: gap of 100 bp
*    20428    21297: contig of 870 bp in length
*    21298    21397: gap of 100 bp
*    21398    22055: contig of 658 bp in length
*    22056    22155: gap of 100 bp
*    22156    23005: contig of 850 bp in length
*    23006    23105: gap of 100 bp
*    23106    24016: contig of 911 bp in length
*    24017    24116: gap of 100 bp
*    24117    24972: contig of 856 bp in length
*    24973    25072: gap of 100 bp
*    25073    25977: contig of 905 bp in length
*    25978    26077: gap of 100 bp
*    26078    26952: contig of 875 bp in length
*    26953    27052: gap of 100 bp
*    27053    27911: contig of 859 bp in length
*    27912    28011: gap of 100 bp
*    28012    28868: contig of 857 bp in length
*    28869    28968: gap of 100 bp
*    28969    29624: contig of 656 bp in length
*    29625    29724: gap of 100 bp
*    29725    30574: contig of 850 bp in length
*    30575    30674: gap of 100 bp
*    30675    31530: contig of 856 bp in length
*    31531    31630: gap of 100 bp
*    31631    32491: contig of 861 bp in length
*    32492    32591: gap of 100 bp
*    32592    33467: contig of 876 bp in length
*    33468    33567: gap of 100 bp
*    33568    34437: contig of 870 bp in length
*    34438    34537: gap of 100 bp
*    34538    35404: contig of 867 bp in length
*    35405    35504: gap of 100 bp
*    35505    36368: contig of 864 bp in length
*    36369    36468: gap of 100 bp
*    36469    37353: contig of 885 bp in length
*    37354    37453: gap of 100 bp
*    37454    38310: contig of 857 bp in length
*    38311    38410: gap of 100 bp
*    38411    39267: contig of 857 bp in length
*    39268    39367: gap of 100 bp
*    39368    40217: contig of 850 bp in length
*    40218    40317: gap of 100 bp
*    40318    41226: contig of 909 bp in length
*    41227    41326: gap of 100 bp
*    41327    42195: contig of 869 bp in length
*    42196    42295: gap of 100 bp
*    42296    43180: contig of 885 bp in length
*    43181    43280: gap of 100 bp
*    43281    44143: contig of 863 bp in length
*    44144    44243: gap of 100 bp
*    44244    45131: contig of 888 bp in length
*    45132    45231: gap of 100 bp
*    45232    46149: contig of 918 bp in length
*    46150    46249: gap of 100 bp
*    46250    47122: contig of 873 bp in length
*    47123    47222: gap of 100 bp
*    47223    48101: contig of 879 bp in length
*    48102    48201: gap of 100 bp
*    48202    49041: contig of 840 bp in length
*    49042    49141: gap of 100 bp
*    49142    50059: contig of 918 bp in length
*    50060    50159: gap of 100 bp
*    50160    51029: contig of 870 bp in length
*    51030    51129: gap of 100 bp
*    51130    51995: contig of 866 bp in length
*    51996    52095: gap of 100 bp
*    52096    52971: contig of 876 bp in length
*    52972    53071: gap of 100 bp
*    53072    53942: contig of 871 bp in length
*    53943    54042: gap of 100 bp
*    54043    54923: contig of 881 bp in length
*    54924    55023: gap of 100 bp
*    55024    55877: contig of 854 bp in length
*    55878    55977: gap of 100 bp
*    55978    56851: contig of 874 bp in length
*    56852    56951: gap of 100 bp
*    56952    57808: contig of 857 bp in length
*    57809    57908: gap of 100 bp
*    57909    58755: contig of 847 bp in length
*    58756    58855: gap of 100 bp
*    58856    59709: contig of 854 bp in length
*    59710    59809: gap of 100 bp
*    59810    60521: contig of 712 bp in length
*    60522    60621: gap of 100 bp
*    60622    61501: contig of 880 bp in length
*    61502    61601: gap of 100 bp
*    61602    62490: contig of 889 bp in length
*    62491    62590: gap of 100 bp
*    62591    63450: contig of 860 bp in length
*    63451    63550: gap of 100 bp
*    63551    64094: contig of 544 bp in length
*    64095    64194: gap of 100 bp
*    64195    65022: contig of 828 bp in length
```

```
            *    65023    65122: gap of 100 bp
            *    65123    65983: contig of 861 bp in length
            *    65984    66083: gap of 100 bp
            *    66084    66922: contig of 839 bp in length
            *    66923    67022: gap of 100 bp
            *    67023    67908: contig of 886 bp in length
            *    67909    68008: gap of 100 bp
            *    68009    68879: contig of 871 bp in length
            *    68880    68979: gap of 100 bp

Alignment Scores:
Pred. No.:              1.21e-48        Length:         118540
Score:                  63.00           Matches:        63
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            4.41%           Indels:         0
DB:                     2               Gaps:           0

US-09-787-097-12 (1-1429) x AC013324 (1-118540)

Qy        921 ProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGly 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      78069 CCAGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGA 78128

Qy        941 AspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAsp 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      78129 GATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGAC 78188

Qy        961 SerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSer 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      78189 TCCAATGCCTACGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGC 78248

Qy        981 ThrCysPro 983
              |||||||||
Db      78249 ACCTGCCCC 78257

RESULT 52
H009384S09
LOCUS       H009384S09              3038 bp    DNA     linear   PRI 14-DEC-2000
DEFINITION  Homo sapiens attractin precursor (ATRN) gene, exons 9 and 10.
ACCESSION   AF218897
VERSION     AF218897.1  GI:8118063
KEYWORDS    .
SEGMENT     9 of 27
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3038)
  AUTHORS   Tang,W., Gunn,T.M., McLaughlin,D.F., Barsh,G.S., Schlossman,S.F.
            and Duke-Cohan,J.S.
  TITLE     Secreted and membrane attractin result from alternative splicing of
            the human ATRN gene
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
  MEDLINE   20283944
   PUBMED   10811918
REFERENCE   2  (bases 1 to 3038)
  AUTHORS   Tang,W., Gunn,T.M., McLaughlin,D.F., Freeman,G.J., Barsh,G.S.,
            Schlossman,S.F. and Duke-Cohan,J.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
            Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
            USA
FEATURES             Location/Qualifiers
     source          1. .3038
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /map="20p13"
     exon            1114. .1297
                     /gene="ATRN"
                     /number=9
     exon            2047. .2201
                     /gene="ATRN"
                     /number=10
ORIGIN

Alignment Scores:
Pred. No.:              5.9e-49         Length:         3038
Score:                  62.00           Matches:        62
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            4.34%           Indels:         0
DB:                     9               Gaps:           0

US-09-787-097-12 (1-1429) x H009384S09 (1-3038)

Qy        483 AspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGly 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1113 GATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGC 1172

Qy        503 HisSerSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAla 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1173 CATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCT 1232

Qy        523 PheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGln 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1233 TTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAG 1292

Qy        543 MetTrp 544
              ||||||
Db       1293 ATGTGG 1298

RESULT 53
H009384S06
LOCUS       H009384S06              1032 bp    DNA     linear   PRI 14-DEC-2000
DEFINITION  Homo sapiens attractin precursor (ATRN) gene, exon 6.
ACCESSION   AF218894
VERSION     AF218894.1  GI:8118060
KEYWORDS    .
SEGMENT     6 of 27
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1032)
  AUTHORS   Tang,W., Gunn,T.M., McLaughlin,D.F., Barsh,G.S., Schlossman,S.F.
            and Duke-Cohan,J.S.
  TITLE     Secreted and membrane attractin result from alternative splicing of
            the human ATRN gene
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
  MEDLINE   20283944
   PUBMED   10811918
REFERENCE   2  (bases 1 to 1032)
  AUTHORS   Tang,W., Gunn,T.M., McLaughlin,D.F., Freeman,G.J., Barsh,G.S.,
            Schlossman,S.F. and Duke-Cohan,J.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
            Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
            USA
FEATURES             Location/Qualifiers
     source          1. .1032
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /map="20p13"
     exon            529. .697
                     /gene="ATRN"
                     /number=6
ORIGIN

Alignment Scores:
Pred. No.:              1.62e-44        Length:         1032
Score:                  57.00           Matches:        57
```

```
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            3.99%          Indels:        0
DB:                     9              Gaps:          0

US-09-787-097-12 (1-1429) x H009384S06 (1-1032)

Qy        315 GlyProGlyCysSerValProValProAlaAsnGlnSerPheTrpThrArgGluGluTyr 334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        528 GGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATAT 587

Qy        335 SerAsnLeuLysLeuProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrp 354
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        588 TCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGG 647

Qy        355 ValValGlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAla 371
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        648 GTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCG 698

RESULT 54
H009384S21
LOCUS       H009384S21              1976 bp    DNA     linear   PRI 14-DEC-2000
DEFINITION  Homo sapiens attractin precursor (ATRN) gene, exon 24.
ACCESSION   AF218909
VERSION     AF218909.1  GI:8118075
KEYWORDS    .
SEGMENT     21 of 27
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1976)
  AUTHORS   Tang,W., Gunn,T.M., McLaughlin,D.F., Barsh,G.S., Schlossman,S.F.
            and Duke-Cohan,J.S.
  TITLE     Secreted and membrane attractin result from alternative splicing of
            the human ATRN gene
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
  MEDLINE   20283944
   PUBMED   10811918
REFERENCE   2  (bases 1 to 1976)
  AUTHORS   Tang,W., Gunn,T.M., McLaughlin,D.F., Freeman,G.J., Barsh,G.S.,
            Schlossman,S.F. and Duke-Cohan,J.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
            Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
            USA
FEATURES             Location/Qualifiers
     source          1. .1976
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /map="20p13"
     exon            392. .549
                     /gene="ATRN"
                     /number=24
ORIGIN

Alignment Scores:
Pred. No.:              1.93e-40       Length:        1976
Score:                  53.00          Matches:       53
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            3.71%          Indels:        0
DB:                     9              Gaps:          0

US-09-787-097-12 (1-1429) x H009384S21 (1-1976)

Qy       1215 AlaGlyThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyr 1234
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 GCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTAC 450

Qy       1235 LysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePhe 1254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 AAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTT 510

Qy       1255 ValTyrValSerAsnPheThrTrpProIleLysIleGln 1267
              |||||||||||||||||||||||||||||||||||||||
Db        511 GTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAG 549

RESULT 55
AB062914S2
LOCUS       AB062914S2              4862 bp    DNA     linear   ROD 16-JAN-2002
DEFINITION  Mesocricetus auratus Atrn gene for attractin, exon 23, 24.
ACCESSION   AB062915
VERSION     AB062915.1  GI:16930137
KEYWORDS    .
SEGMENT     2 of 5
SOURCE      Mesocricetus auratus (golden hamster)
  ORGANISM  Mesocricetus auratus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
            Mesocricetus.
REFERENCE   1
  AUTHORS   Kuramoto,T., Nomoto,T., Fujiwara,A., Mizutani,M., Sugimura,T. and
            Ushijima,T.
  TITLE     Insertional mutation of the Attractin gene in the black tremor
            hamster
  JOURNAL   Mamm. Genome 13 (1), 36-40 (2002)
  MEDLINE   21635551
   PUBMED   11773967
REFERENCE   2  (bases 1 to 4862)
  AUTHORS   Kuramoto,T. and Ushijima,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-JUN-2001) Takashi Kuramoto, National Cancer Center
            Research Institute, Carcinogenesis Division; Tsukiji 5-1-1,
            Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramot@gan2.ncc.go.jp,
            URL:http://www.ncc.go.jp/research/rat-genome/, Tel:81-3-3542-2511,
            Fax:81-3-5565-1753)
FEATURES             Location/Qualifiers
     source          1. .4862
                     /organism="Mesocricetus auratus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10036"
     exon            594. .672
                     /gene="Atrn"
                     /product="attractin"
                     /note="CDS is reported in Acc# AB062918"
                     /number=23
     exon            2976. .3133
                     /gene="Atrn"
                     /product="attractin"
                     /note="CDS is reported in Acc# AB062918"
                     /number=24
     gene            3782. .3886
                     /gene="Atrn"
     repeat_region   3782. .3886
                     /rpt_type=tandem
                     /rpt_unit="aag"
ORIGIN

Alignment Scores:
Pred. No.:              3.97e-40       Length:        4862
Score:                  53.00          Matches:       53
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            3.71%          Indels:        0
DB:                     10             Gaps:          0

US-09-787-097-12 (1-1429) x AB062914S2 (1-4862)

Qy       1215 AlaGlyThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyr 1234
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2975 GCTGGAACCCAGGCAGGAGAAGAAATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTAC 3034

Qy       1235 LysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePhe 1254
```

```
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3035 AAAGATAGCTTCTCTAATGAGAAATTTGATTTTCGCAACCATCCAAACATTACTTTCTTT 3094

Qy       1255 ValTyrValSerAsnPheThrTrpProIleLysIleGln 1267
                    |||||||||||||||||||||||||||||||||||||
Db       3095 GTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAG 3133

RESULT 56
AC055771
LOCUS       AC055771               167357 bp    DNA     linear   HTG 24-AUG-2002
DEFINITION  Homo sapiens chromosome 20 clone RP11-561P16 map 20, WORKING DRAFT
            SEQUENCE, 21 unordered pieces.
ACCESSION   AC055771
VERSION     AC055771.2  GI:7770449
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 167357)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 20, clone RP11-561P16
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 167357)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-APR-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 167357)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On May 12, 2000 this sequence version replaced gi:7582593.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
               Center: Whitehead Institute/ MIT Center for Genome Research
               Center code: WIBR
               Web site: http://www-seq.wi.mit.edu
               Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
               Center project name: L5007
               Center clone name: 561_P_16
            -------------- Summary Statistics
               Sequencing vector: M13; M77815; 100% of reads
               Chemistry: Dye-terminator Big Dye; 100% of reads
               Assembly program: Phrap; version 0.960731
               Consensus quality: 156504 bases at least Q40
               Consensus quality: 161679 bases at least Q30
               Consensus quality: 163964 bases at least Q20
               Insert size: 170000; agarose-fp
               Insert size: 165357; sum-of-contigs
               Quality coverage: 4.5 in Q20 bases; agarose-fp
               Quality coverage: 4.6 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 21 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1216: contig of 1216 bp in length
            *     1217     1316: gap of 100 bp
            *     1317     2985: contig of 1669 bp in length
            *     2986     3085: gap of 100 bp
            *     3086     4628: contig of 1543 bp in length
            *     4629     4728: gap of 100 bp
            *     4729     7507: contig of 2779 bp in length
            *     7508     7607: gap of 100 bp
            *     7608    11292: contig of 3685 bp in length
            *    11293    11392: gap of 100 bp
            *    11393    14213: contig of 2821 bp in length
            *    14214    14313: gap of 100 bp
            *    14314    17828: contig of 3515 bp in length
            *    17829    17928: gap of 100 bp
            *    17929    21068: contig of 3140 bp in length
            *    21069    21168: gap of 100 bp
            *    21169    25035: contig of 3867 bp in length
            *    25036    25135: gap of 100 bp
            *    25136    28331: contig of 3196 bp in length
            *    28332    28431: gap of 100 bp
            *    28432    33595: contig of 5164 bp in length
            *    33596    33695: gap of 100 bp
            *    33696    40346: contig of 6651 bp in length
            *    40347    40446: gap of 100 bp
            *    40447    43807: contig of 3361 bp in length
            *    43808    43907: gap of 100 bp
            *    43908    48727: contig of 4820 bp in length
            *    48728    48827: gap of 100 bp
            *    48828    56083: contig of 7256 bp in length
            *    56084    56183: gap of 100 bp
            *    56184    67065: contig of 10882 bp in length
            *    67066    67165: gap of 100 bp
            *    67166    82881: contig of 15716 bp in length
            *    82882    82981: gap of 100 bp
            *    82982   101845: contig of 18864 bp in length
            *   101846   101945: gap of 100 bp
            *   101946   122755: contig of 20810 bp in length
            *   122756   122855: gap of 100 bp
            *   122856   142457: contig of 19602 bp in length
            *   142458   142557: gap of 100 bp
            *   142558   167357: contig of 24800 bp in length.
```

```
FEATURES             Location/Qualifiers
     source          1. .167357
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /map="20"
                     /clone="RP11-561P16"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1216
                     /note="assembly_fragment"
     misc_feature    1317. .2985
                     /note="assembly_fragment"
     misc_feature    3086. .4628
                     /note="assembly_fragment"
     misc_feature    4729. .7507
                     /note="assembly_fragment"
     misc_feature    7608. .11292
                     /note="assembly_fragment"
     misc_feature    11393. .14213
                     /note="assembly_fragment"
     misc_feature    14314. .17828
                     /note="assembly_fragment"
     misc_feature    17929. .21068
                     /note="assembly_fragment"
     misc_feature    21169. .25035
                     /note="assembly_fragment"
     misc_feature    25136. .28331
                     /note="assembly_fragment"
     misc_feature    28432. .33595
                     /note="assembly_fragment"
     misc_feature    33696. .40346
                     /note="assembly_fragment"
     misc_feature    40447. .43807
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:right"
     misc_feature    43908. .48727
                     /note="assembly_fragment"
     misc_feature    48828. .56083
                     /note="assembly_fragment"
     misc_feature    56184. .67065
                     /note="assembly_fragment"
     misc_feature    67166. .82881
                     /note="assembly_fragment"
     misc_feature    82982. .101845
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
     misc_feature    101946. .122755
                     /note="assembly_fragment"
     misc_feature    122856. .142457
                     /note="assembly_fragment"
     misc_feature    142558. .167357
                     /note="assembly_fragment"
ORIGIN

Alignment Scores:
Pred. No.:              6.77e-39        Length:         167357
Score:                  53.00           Matches:        53
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            3.71%           Indels:         0
DB:                     2               Gaps:           0

US-09-787-097-12 (1-1429) x AC055771 (1-167357)

Qy       1215 AlaGlyThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyr 1234
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2354 GCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTAC 2413

Qy       1235 LysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePhe 1254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       2414 AAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTT 2473

Qy       1255 ValTyrValSerAsnPheThrTrpProIleLysIleGln 1267
              |||||||||||||||||||||||||||||||||||||||
Db       2474 GTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAG 2512

RESULT 57
AC017113
LOCUS       AC017113               206329 bp    DNA     linear   HTG 07-JUL-2000
DEFINITION  Homo sapiens chromosome 20 clone RP11-574H7, WORKING DRAFT
            SEQUENCE, 31 unordered pieces.
ACCESSION   AC017113
VERSION     AC017113.4  GI:8570339
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 206329)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 206329)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-DEC-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Jun 17, 2000 this sequence version replaced gi:6855220.

            ---------------- Genome Center ----------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            -------------- Project Information --------------
            Center project name: H_NH0574H07
            -------------- Summary Statistics  --------------
            Sequencing vector: M13; 91%
            Sequencing vector: plasmid; 9%
            Chemistry: Dye-primer ET; 91% of reads
            Chemistry: Dye-terminator Big Dye; 9% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 184921 bases at least Q40
            Consensus quality: 192397 bases at least Q30
            Consensus quality: 195862 bases at least Q20
            Insert size: 237000; agarose-fp
            Insert size: 203329; sum-of-contigs
            Quality coverage: 3.84 in Q20 bases; agarose-fp
            Quality coverage: 3.66 in Q20 bases; sum-of-contigs
            ------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 31 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1539: contig of 1539 bp in length
            *     1540     1639: gap of unknown length
            *     1640     3259: contig of 1620 bp in length
            *     3260     3359: gap of unknown length
            *     3360     5400: contig of 2041 bp in length
            *     5401     5500: gap of unknown length
            *     5501     7783: contig of 2283 bp in length
            *     7784     7883: gap of unknown length
            *     7884    11241: contig of 3358 bp in length
            *    11242    11341: gap of unknown length
            *    11342    13482: contig of 2141 bp in length
            *    13483    13582: gap of unknown length
            *    13583    16915: contig of 3333 bp in length
            *    16916    17015: gap of unknown length
            *    17016    21346: contig of 4331 bp in length
```

```
            *    21347    21446: gap of unknown length
            *    21447    25103: contig of 3657 bp in length
            *    25104    25203: gap of unknown length
            *    25204    28415: contig of 3212 bp in length
            *    28416    28515: gap of unknown length
            *    28516    32873: contig of 4358 bp in length
            *    32874    32973: gap of unknown length
            *    32974    36496: contig of 3523 bp in length
            *    36497    36596: gap of unknown length
            *    36597    40649: contig of 4053 bp in length
            *    40650    40749: gap of unknown length
            *    40750    44878: contig of 4129 bp in length
            *    44879    44978: gap of unknown length
            *    44979    50353: contig of 5375 bp in length
            *    50354    50453: gap of unknown length
            *    50454    55946: contig of 5493 bp in length
            *    55947    56046: gap of unknown length
            *    56047    61420: contig of 5374 bp in length
            *    61421    61520: gap of unknown length
            *    61521    67819: contig of 6299 bp in length
            *    67820    67919: gap of unknown length
            *    67920    72327: contig of 4408 bp in length
            *    72328    72427: gap of unknown length
            *    72428    77293: contig of 4866 bp in length
            *    77294    77393: gap of unknown length
            *    77394    82626: contig of 5233 bp in length
            *    82627    82726: gap of unknown length
            *    82727    90186: contig of 7460 bp in length
            *    90187    90286: gap of unknown length
            *    90287    96584: contig of 6298 bp in length
            *    96585    96684: gap of unknown length
            *    96685   106019: contig of 9335 bp in length
            *   106020   106119: gap of unknown length
            *   106120   113530: contig of 7411 bp in length
            *   113531   113630: gap of unknown length
            *   113631   125151: contig of 11521 bp in length
            *   125152   125251: gap of unknown length
            *   125252   137677: contig of 12426 bp in length
            *   137678   137777: gap of unknown length
            *   137778   151263: contig of 13486 bp in length
            *   151264   151363: gap of unknown length
            *   151364   165370: contig of 14007 bp in length
            *   165371   165470: gap of unknown length
            *   165471   184704: contig of 19234 bp in length
            *   184705   184804: gap of unknown length
            *   184805   206329: contig of 21525 bp in length.
FEATURES             Location/Qualifiers
     source          1. .206329
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /clone="RP11-574H7"

ORIGIN

Alignment Scores:
Pred. No.:              8.01e-39        Length:         206329
Score:                  53.00           Matches:        53
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            3.71%           Indels:         0
DB:                     2               Gaps:           0

US-09-787-097-12 (1-1429) x AC017113 (1-206329)

Qy       1215 AlaGlyThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyr 1234
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     173249 GCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTAC 173308

Qy       1235 LysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePhe 1254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     173309 AAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTT 173368

Qy       1255 ValTyrValSerAsnPheThrTrpProIleLysIleGln 1267
              |||||||||||||||||||||||||||||||||||||||
Db     173369 GTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAG 173407

RESULT 58
AR164811/c
LOCUS       AR164811                226 bp    DNA     linear   PAT 17-OCT-2001
DEFINITION  Sequence 6 from patent US 6274339.
ACCESSION   AR164811
VERSION     AR164811.1  GI:16238018
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 226)
  AUTHORS   Moore,K. and Nagle,D.Lynn.
  TITLE     Methods and compositions for the diagnosis and treatment of body
            weight disorders, including obesity
  JOURNAL   Patent: US 6274339-A 6 14-AUG-2001;
FEATURES             Location/Qualifiers
     source          1. .226
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:              2.86e-39        Length:         226
Score:                  51.00           Matches:        51
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            3.57%           Indels:         0
DB:                     6               Gaps:           0

US-09-787-097-12 (1-1429) x AR164811 (1-226)

Qy        871 AspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSerAspAla 890
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        225 GATATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGTGATGCT 166

Qy        891 GlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThrCysIle 910
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        165 GGCTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGATTAAAGGCTGCAACCTGCATC 106

Qy        911 AsnProLeuAsnGlySerValCysGluArgPro 921
              |||||||||||||||||||||||||||||||
Db        105 AACCCTCTCAATGGCAGCGTCTGTGAAAGGCCT 73

RESULT 59
AB06968S09
LOCUS       AB06968S09              1050 bp    DNA     linear   ROD 04-APR-2002
DEFINITION  Rattus norvegicus Atrn gene for membrane attractin, exon 10.
ACCESSION   AB049230
VERSION     AB049230.1  GI:12275371
KEYWORDS    .
SEGMENT     9 of 27
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1
  AUTHORS   Kuramoto,T., Kitada,K., Inui,T., Sasaki,Y., Ito,K., Hase,T.,
            Kawaguchi,S., Ogawa,Y., Nakao,K., Barsh,G.S., Nagao,M., Ushijima,T.
            and Serikawa,T.
  TITLE     Attractin/mahogany/zitter plays a critical role in myelination of
            the central nervous system
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (2), 559-564 (2001)
  MEDLINE   21143347
   PUBMED   11209055
REFERENCE   2  (bases 1 to 1050)
  AUTHORS   Kuramoto,T. and Ushijima,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-SEP-2000) Takashi Kuramoto, National Cancer Center
```

```
           Research Institute, Carcinogenesis Division; Tsukiji 5-1-1,
           Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramot@gan2.ncc.go.jp,
           Tel:81-33542-2511, Fax:81-5565-1753)
FEATURES             Location/Qualifiers
     source          1. .1050
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /chromosome="3"
     exon            496. .650
                     /gene="Atrn"
                     /product="membrane attractin"
                     /note="CDS is reported in Acc# AB049248"
                     /number=10
ORIGIN

Alignment Scores:
Pred. No.:              9.8e-39       Length:        1050
Score:                  51.00         Matches:       51
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            3.57%         Indels:        0
DB:                     10            Gaps:          0

US-09-787-097-12 (1-1429) x AB06968S09 (1-1050)

Qy        545 ThrIleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerGly 564
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 ACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCATACAGCTGTGATAGTGAGTGGA 556

Qy        565 ThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLys 584
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 ACCATGCTGGTGTTCGGAGGGAACACACACAATGACACTTCCATGAGCCACGGTGCCAAA 616

Qy        585 CysPheSerSerAspPheMetAlaTyrAspIle 595
              |||||||||||||||||||||||||||||||
Db        617 TGCTTCTCCTCAGACTTCATGGCTTATGACATT 649

RESULT 60
HS741H3S
LOCUS       HS741H3S                 553 bp    DNA     linear   STS 24-AUG-1999
DEFINITION  H.sapiens STS from genomic clone 741H3, sequence tagged site.
ACCESSION   AL110053
VERSION     AL110053.1  GI:5777509
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 553)
  AUTHORS   Deloukas,P., O'Neill,L., Holden,J., Mistry,D., Huckle,E., Taylor,R.
            and Hunt,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-AUG-1999) E-mail contact: humquery@sanger.ac.uk
COMMENT     Marker stdJ741H3SP6 (Primer A : AGGTGAAGCATGTGACATTCC; Primer B :
            AAAGAAGCTCCTACCCTGCC; amplimer size : 124 bp) is from sequence
            generated from the SP6 end of PAC 741H3. 741H3 is part of the
            bacterial clone contigs constructed by the Chromosome 20 Mapping
            Group.(http://www.sanger.ac.uk/HGP/Chr20/) 741H3 is from the
            library  constructed at the Roswell Park Cancer Institute by the
            group of Pieter de Jong.
            For further details see http://bacpac.med.buffalo.edu/.
FEATURES             Location/Qualifiers
     source          1. .553
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /clone="741H3"
ORIGIN

Alignment Scores:
Pred. No.:              4.53e-36      Length:        553
Score:                  48.00         Matches:       48
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            3.36%         Indels:        0
DB:                     11            Gaps:          0

US-09-787-097-12 (1-1429) x HS741H3S (1-553)

Qy        268 ThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIleProHisCys 287
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 ACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGT 63

Qy        288 ThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspValArgGlyCys 307
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 ACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGC 123

Qy        308 SerCysPheSerAspTrpGlnGly 315
              |||||||||||||||||||||||||
Db        124 TCCTGCTTCTCAGACTGGCAGGGT 147

RESULT 61
H009384S05
LOCUS       H009384S05               1232 bp    DNA     linear   PRI 14-DEC-2000
DEFINITION  Homo sapiens attractin precursor (ATRN) gene, exon 5.
ACCESSION   AF218893
VERSION     AF218893.1  GI:8118059
KEYWORDS    .
SEGMENT     5 of 27
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1232)
  AUTHORS   Tang,W., Gunn,T.M., McLaughlin,D.F., Barsh,G.S., Schlossman,S.F.
            and Duke-Cohan,J.S.
  TITLE     Secreted and membrane attractin result from alternative splicing of
            the human ATRN gene
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
  MEDLINE   20283944
   PUBMED   10811918
REFERENCE   2  (bases 1 to 1232)
  AUTHORS   Tang,W., Gunn,T.M., McLaughlin,D.F., Freeman,G.J., Barsh,G.S.,
            Schlossman,S.F. and Duke-Cohan,J.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
            Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
            USA
FEATURES             Location/Qualifiers
     source          1. .1232
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /map="20p13"
     exon            621. .826
                     /gene="ATRN"
                     /number=5
ORIGIN

Alignment Scores:
Pred. No.:              8.61e-36      Length:        1232
Score:                  48.00         Matches:       48
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            3.36%         Indels:        0
DB:                     9             Gaps:          0

US-09-787-097-12 (1-1429) x H009384S05 (1-1232)

Qy        268 ThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIleProHisCys 287
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        685 ACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGT 744
```

```
Qy        288 ThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspValArgGlyCys 307
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        745 ACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGGATGC 804

Qy        308 SerCysPheSerAspTrpGlnGly 315
              |||||||||||||||||||||||||
Db        805 TCCTGCTTCTCAGACTGGCAGGGT 828

RESULT 62
H009384S16
LOCUS       H009384S16              1674 bp    DNA     linear   PRI 14-DEC-2000
DEFINITION  Homo sapiens attractin precursor (ATRN) gene, exon 19.
ACCESSION   AF218904
VERSION     AF218904.1  GI:8118070
KEYWORDS    .
SEGMENT     16 of 27
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1674)
  AUTHORS   Tang,W., Gunn,T.M., McLaughlin,D.F., Barsh,G.S., Schlossman,S.F.
            and Duke-Cohan,J.S.
  TITLE     Secreted and membrane attractin result from alternative splicing of
            the human ATRN gene
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
  MEDLINE   20283944
   PUBMED   10811918
REFERENCE   2  (bases 1 to 1674)
  AUTHORS   Tang,W., Gunn,T.M., McLaughlin,D.F., Freeman,G.J., Barsh,G.S.,
            Schlossman,S.F. and Duke-Cohan,J.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
            Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
            USA
FEATURES             Location/Qualifiers
     source          1. .1674
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /map="20p13"
     exon            810. .947
                     /gene="ATRN"
                     /number=19
ORIGIN

Alignment Scores:
Pred. No.:                9.27e-34        Length:         1674
Score:                    46.00           Matches:        46
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              3.22%           Indels:         0
DB:                       9               Gaps:           0

US-09-787-097-12 (1-1429) x H009384S16 (1-1674)

Qy       1062 AlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGlu 1081
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        809 GCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAG 868

Qy       1082 AsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThr 1101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        869 AACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACC 928

Qy       1102 AsnGlyGlyLysCysGln 1107
              |||||||||||||||||
Db        929 AATGGAGGGAAATGTCAG 946

RESULT 63
AR164812/c
LOCUS       AR164812                221 bp    DNA     linear   PAT 17-OCT-2001
DEFINITION  Sequence 7 from patent US 6274339.
ACCESSION   AR164812
VERSION     AR164812.1  GI:16238020
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 221)
  AUTHORS   Moore,K. and Nagle,D.Lynn.
  TITLE     Methods and compositions for the diagnosis and treatment of body
            weight disorders, including obesity
  JOURNAL   Patent: US 6274339-A 7 14-AUG-2001;
FEATURES             Location/Qualifiers
     source          1. .221
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:                1.68e-33        Length:         221
Score:                    45.00           Matches:        45
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              3.15%           Indels:         0
DB:                       6               Gaps:           0

US-09-787-097-12 (1-1429) x AR164812 (1-221)

Qy        871 AspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSerAspAla 890
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 GATATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGTGATGCT 161

Qy        891 GlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThrCysIle 910
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 GGCTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGATTAAAGGCTGCAACCTGCATC 101

Qy        911 AsnProLeuAsnGly 915
              |||||||||||||||
Db        100 AACCCTCTCAATGGT 86

RESULT 64
AB06968S16
LOCUS       AB06968S16              1880 bp    DNA     linear   ROD 04-APR-2002
DEFINITION  Rattus norvegicus Atrn gene for membrane attractin, exon 17, 18.
ACCESSION   AB049237
VERSION     AB049237.1  GI:12275378
KEYWORDS    .
SEGMENT     16 of 27
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1
  AUTHORS   Kuramoto,T., Kitada,K., Inui,T., Sasaki,Y., Ito,K., Hase,T.,
            Kawaguchi,S., Ogawa,Y., Nakao,K., Barsh,G.S., Nagao,M., Ushijima,T.
            and Serikawa,T.
  TITLE     Attractin/mahogany/zitter plays a critical role in myelination of
            the central nervous system
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (2), 559-564 (2001)
  MEDLINE   21143347
   PUBMED   11209055
REFERENCE   2  (bases 1 to 1880)
  AUTHORS   Kuramoto,T. and Ushijima,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-SEP-2000) Takashi Kuramoto, National Cancer Center
            Research Institute, Carcinogenesis Division; Tsukiji 5-1-1,
            Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramot@gan2.ncc.go.jp,
            Tel:81-33542-2511, Fax:81-5565-1753)
FEATURES             Location/Qualifiers
     source          1. .1880
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
```

```
                /chromosome="3"
     exon       595. .780
                /gene="Atrn"
                /product="membrane attractin"
                /note="CDS is reported in Acc# AB049248"
                /number=17
     exon       1283. .1516
                /gene="Atrn"
                /product="membrane attractin"
                /note="CDS is reported in Acc# AB049248"
                /number=18
ORIGIN

Alignment Scores:
Pred. No.:              9.34e-33        Length:         1880
Score:                  45.00           Matches:        45
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            3.15%           Indels:         0
DB:                     10              Gaps:           0

US-09-787-097-12 (1-1429) x AB06968S16 (1-1880)

Qy        985 GluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrp 1004
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1285 GAAAATTGCTCTGGCTACTGTACCTGCAGCCATTGCTTGGAGCAGCCAGGCTGTGGTTGG 1344

Qy       1005 CysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyPro 1024
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1345 TGTACTGATCCTAGCAATACTGGGAAAGGCAAATGTATTGAGGGAAGCTATAAAGGACCC 1404

Qy       1025 ValLysMetProSer 1029
              ||||||||||||||
Db       1405 GTGAAGATGCCTTCA 1419

RESULT 65
H009384S04
LOCUS       H009384S04              1152 bp    DNA     linear   PRI 14-DEC-2000
DEFINITION  Homo sapiens attractin precursor (ATRN) gene, exon 4.
ACCESSION   AF218892
VERSION     AF218892.1  GI:8118058
KEYWORDS    .
SEGMENT     4 of 27
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1152)
  AUTHORS   Tang,W., Gunn,T.M., McLaughlin,D.F., Barsh,G.S., Schlossman,S.F.
            and Duke-Cohan,J.S.
  TITLE     Secreted and membrane attractin result from alternative splicing of
            the human ATRN gene
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
  MEDLINE   20283944
   PUBMED   10811918
REFERENCE   2  (bases 1 to 1152)
  AUTHORS   Tang,W., Gunn,T.M., McLaughlin,D.F., Freeman,G.J., Barsh,G.S.,
            Schlossman,S.F. and Duke-Cohan,J.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
            Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
            USA
FEATURES             Location/Qualifiers
     source          1. .1152
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /map="20p13"
     exon            818. .946
                     /gene="ATRN"
                     /number=4
ORIGIN

Alignment Scores:
Pred. No.:              5.79e-32        Length:         1152
Score:                  44.00           Matches:        44
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            3.08%           Indels:         0
DB:                     9               Gaps:           0

US-09-787-097-12 (1-1429) x H009384S04 (1-1152)

Qy        202 PheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGluValValAla 221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        813 TTCAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCC 872

Qy        222 ThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPhe 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        873 ACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGATTT 932

Qy        242 AsnIleThrTyr 245
              ||||||||||||
Db        933 AATATTACTTAC 944

RESULT 66
AB06968S22
LOCUS       AB06968S22              1474 bp    DNA     linear   ROD 04-APR-2002
DEFINITION  Rattus norvegicus Atrn gene for membrane attractin, exon 24.
ACCESSION   AB049243
VERSION     AB049243.1  GI:12275384
KEYWORDS    .
SEGMENT     22 of 27
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1
  AUTHORS   Kuramoto,T., Kitada,K., Inui,T., Sasaki,Y., Ito,K., Hase,T.,
            Kawaguchi,S., Ogawa,Y., Nakao,K., Barsh,G.S., Nagao,M., Ushijima,T.
            and Serikawa,T.
  TITLE     Attractin/mahogany/zitter plays a critical role in myelination of
            the central nervous system
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (2), 559-564 (2001)
  MEDLINE   21143347
   PUBMED   11209055
REFERENCE   2  (bases 1 to 1474)
  AUTHORS   Kuramoto,T. and Ushijima,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-SEP-2000) Takashi Kuramoto, National Cancer Center
            Research Institute, Carcinogenesis Division; Tsukiji 5-1-1,
            Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramot@gan2.ncc.go.jp,
            Tel:81-33542-2511, Fax:81-5565-1753)
FEATURES             Location/Qualifiers
     source          1. .1474
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /chromosome="3"
     exon            366. .523
                     /gene="Atrn"
                     /product="membrane attractin"
                     /note="CDS is reported in Acc# AB049248"
                     /number=24
ORIGIN

Alignment Scores:
Pred. No.:              7.05e-32        Length:         1474
Score:                  44.00           Matches:        44
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            3.08%           Indels:         0
DB:                     10              Gaps:           0
```

```
US-09-787-097-12 (1-1429) x AB06968S22 (1-1474)

Qy       1224 ProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPhe 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        392 CCTGTTGTTTCAAAAACCAACATCAAGGAATACAAAGATAGCTTCTCTAATGAGAAATTT 451

Qy       1244 AspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpPro 1263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        452 GATTTTCGCAACCATCCAAACATTACTTTCTTTGTTTATGTCAGTAATTTCACTTGGCCC 511

Qy       1264 IleLysIleGln 1267
              ||||||||||||
Db        512 ATCAAAATTCAA 523

RESULT 67
HSDJ581P3
LOCUS       HSDJ581P3              73666 bp    DNA     linear   PRI 23-FEB-2001
DEFINITION  Human DNA sequence from clone RP4-581P3 on chromosome 20. Contains
            an SF3A3 (splicing factor 3a, subunit 3, 60kD (SAP61)) pseudogene,
            part of the ATRN gene for attractin (with dipeptidylpeptidase IV
            activity), ESTs, STSs and GSSs, complete sequence.
ACCESSION   AL109805
VERSION     AL109805.14  GI:7321191
KEYWORDS    HTG; ATRN; attractin; SAP61; SF3A3.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 73666)
  AUTHORS   Ramsay,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Mar 23, 2000 this sequence version replaced gi:7228123.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 20, constructed by the Sanger Centre Chromosome 20
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr20
            IMPORTANT: This sequence is not the entire insert of clone
            RP4-581P3 It may be shorter because we sequence overlapping
            sections only once, except for a 100 base overlap.
            The true left end of clone RP4-581P3 is at 1 in this sequence. The
            true left end of clone RP4-741H3 is at 73567 in this sequence. This
            sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. RP4-581P3 is from the
            library RPCI-4 constructed by the group of Pieter de Jong. For
            further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pCYPAC2.
FEATURES             Location/Qualifiers
     source          1. .73666
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /clone="RP4-581P3"
                     /clone_lib="RPCI-4"
     repeat_region   61. .361
                     /note="L2 repeat: matches 1703. .2052 of consensus"
     repeat_region   665. .969
                     /note="AluY repeat: matches 1. .306 of consensus"
     repeat_region   990. .1292
                     /note="AluSg repeat: matches 1. .303 of consensus"
     repeat_region   1551. .1626
                     /note="MIR repeat: matches 162. .247 of consensus"
     repeat_region   2105. .2407
                     /note="AluJb repeat: matches 5. .306 of consensus"
     repeat_region   2467. .2599
                     /note="MIR repeat: matches 6. .149 of consensus"
     repeat_region   2817. .3115
                     /note="AluSq repeat: matches 3. .300 of consensus"
     repeat_region   3233. .3275
                     /note="L2 repeat: matches 2617. .2659 of consensus"
     repeat_region   3310. .3352
                     /note="L2 repeat: matches 2617. .2659 of consensus"
     repeat_region   3387. .3427
                     /note="L2 repeat: matches 2617. .2657 of consensus"
     repeat_region   3457. .3506
                     /note="L2 repeat: matches 2570. .2659 of consensus"
     repeat_region   3534. .3583
                     /note="L2 repeat: matches 2570. .2659 of consensus"
     repeat_region   3618. .3660
                     /note="L2 repeat: matches 2617. .2659 of consensus"
     repeat_region   3688. .3737
                     /note="L2 repeat: matches 2570. .2659 of consensus"
     repeat_region   3765. .3814
                     /note="L2 repeat: matches 2570. .2659 of consensus"
     repeat_region   3849. .4008
                     /note="L2 repeat: matches 2617. .2750 of consensus"
     repeat_region   4452. .4924
                     /note="L1PA7 repeat: matches 5664. .6143 of consensus"
     repeat_region   5160. .5428
                     /note="AluJo repeat: matches 1. .304 of consensus"
     repeat_region   5463. .5826
                     /note="L2 repeat: matches 2348. .2706 of consensus"
     repeat_region   5827. .6274
                     /note="MLT1C repeat: matches 1. .466 of consensus"
     repeat_region   6275. .6350
                     /note="L2 repeat: matches 2272. .2348 of consensus"
     misc_feature    6783. .7189
                     /note="match: GSS: Em:AQ585728"
     misc_feature    6786. .6963
                     /note="match: GSS: Em:AQ585784"
     repeat_region   6874. .6912
                     /note="MLT1A1 repeat: matches 323. .361 of consensus"
     repeat_region   7594. .7881
                     /note="AluJo repeat: matches 3. .285 of consensus"
     misc_feature    complement(7666. .8180)
                     /note="match: GSS: Em:AQ453150"
     repeat_region   8025. .8050
                     /note="13 copies 2 mer tt 92% conserved"
     repeat_region   9591. .9884
                     /note="AluJb repeat: matches 1. .294 of consensus"
     repeat_region   10205. .10511
                     /note="AluJo repeat: matches 1. .307 of consensus"
     misc_feature    complement(10997. .11232)
                     /note="match: STS: Em:G08041"
     repeat_region   11393. .11577
                     /note="L1M4 repeat: matches 4132. .4326 of consensus"
     repeat_region   11683. .11848
                     /note="L1 repeat: matches 3736. .3898 of consensus"
     repeat_region   11849. .12157
                     /note="AluSg repeat: matches 1. .309 of consensus"
     repeat_region   12158. .12182
                     /note="L1 repeat: matches 3712. .3736 of consensus"
     repeat_region   12200. .12396
                     /note="L1MEc repeat: matches 1702. .1912 of consensus"
     repeat_region   12417. .12586
                     /note="L1MC5 repeat: matches 7758. .7913 of consensus"
```

```
     repeat_region   12634. .12914
                     /note="AluSx repeat: matches 6. .312 of consensus"
     repeat_region   12958. .13404
                     /note="L1MEc repeat: matches 1267. .1711 of consensus"
     repeat_region   13486. .13733
                     /note="AluSg1 repeat: matches 36. .285 of consensus"
     repeat_region   14313. .14415
                     /note="MER86 repeat: matches 89. .196 of consensus"
     repeat_region   14493. .15186
                     /note="L1 repeat: matches 4409. .5090 of consensus"
     repeat_region   15187. .15487
                     /note="AluSx repeat: matches 12. .312 of consensus"
     repeat_region   15488. .15539
                     /note="L1 repeat: matches 4358. .4409 of consensus"
     misc_feature    15537. .15942
                     /note="match: STS: Em:R15876"
     misc_feature    15537. .15748
                     /note="match: STS: Em:G15134"
     gene            complement(15767. .17243)
                     /gene="dJ581P3.1"
                     /pseudo
     CDS             complement(15767. .17243)
                     /gene="dJ581P3.1"
                     /note="dJ581P3.1 (SF3A3 (splicing factor 3a, subunit 3,
                     60kD (SAP61)) pseudogene)
                     match: cDNAs: Em:X81789 Em:U08815
                     match: ESTs: Em:AI128669 Em:T56818 Em:AW470436 Em:AA300516
                     Em:AA703023 Em:AA760210 Em:AA054241 Em:AA443032 Em:R17811
                     Em:AW493824 Em:H67476 Em:AI020542 Em:AA343902 Em:H89818
                     Em:AA308239 Em:AA311786 Em:AI541046 Em:AA239392
                     Em:AA694306 Em:AA086662 Em:R15877 Em:AI226265 Em:F22705
                     Em:AI155866 Em:C06258 Em:AA086076 Em:AA081903 Em:AI019200
                     Em:AV138764 Em:AW406960 Em:AI202796 Em:R00790 Em:AA308500
                     Em:AA003841 Em:AV290458 Em:AV166565 Em:AA074650
                     Em:AA135375 Em:AA355911 Em:AA730707 Em:AA068453 Em:T27366
                     Em:AA306883 Em:AL035731 Em:AA992803 Em:AL035732
                     Em:AA854771 Em:AW320502 Em:AA599259 Em:AI374964
                     Em:AI338280 Em:AA662262 Em:AA642724 Em:W83741 Em:AA662366
                     Em:AA463475 Em:W35845 Em:AI269312 Em:AL118208 Em:AA986711
                     Em:AI374716 Em:AI908435 Em:C82711 Em:AL039268 Em:AW246145
                     match: proteins: Tr:Q12874 Wp:T13H5 Tr:O59706 Tr:Q22469
                     Tr:O46106 Sw:P19736"
                     /pseudo
                     /codon_start=1
                     /evidence=not_experimental
     misc_feature    16008. .16100
                     /note="match: GSS: Em:AQ279830"
     misc_feature    16104. .16193
                     /note="match: GSS: Em:AQ074332"
     misc_feature    complement(16881. .16954)
                     /gene="dJ581P3.1"
                     /note="match: GSS: Em:AQ975997"
                     /pseudo
     misc_feature    16980. .17146
                     /note="match: GSS: Em:AQ373489"
     misc_feature    17150. .17265
                     /note="match: GSS: Em:AQ373489"
     repeat_region   17389. .17749
                     /note="L2 repeat: matches 1140. .1536 of consensus"
     repeat_region   17859. .18030
                     /note="L2 repeat: matches 1791. .1990 of consensus"
     repeat_region   18405. .18442
                     /note="19 copies 2 mer gt 97% conserved"
     misc_feature    complement(19050. .19224)
                     /note="match: GSS: Em:AQ472547"
     repeat_region   19051. .19092
                     /note="21 copies 2 mer ag 90% conserved"
     misc_feature    complement(19093. .19250)
                     /note="match: GSS: Em:AQ271994"
     repeat_region   19142. .19175

Alignment Scores:
Pred. No.:              1.62e-30        Length:         73666
```

```
Score:                  44.00           Matches:        44
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            3.08%           Indels:         0
DB:                     9               Gaps:           0

US-09-787-097-12 (1-1429) x HSDJ581P3 (1-73666)

Qy        202 PheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGluValValAla 221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      71995 TTCAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCC 72054

Qy        222 ThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPhe 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      72055 ACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGATTT 72114

Qy        242 AsnIleThrTyr 245
              ||||||||||||
Db      72115 AATATTACTTAC 72126

RESULT 68
AF273610
LOCUS       AF273610                237 bp    mRNA    linear   MAM 09-JUL-2000
DEFINITION  Sus scrofa mahogany mRNA, partial cds.
ACCESSION   AF273610
VERSION     AF273610.1  GI:8980838
KEYWORDS    .
SOURCE      Sus scrofa (pig)
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 237)
  AUTHORS   Matteri,R.L. and Dyer,C.J.
  TITLE     Sus scrofa mahogany mRNA
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 237)
  AUTHORS   Matteri,R.L. and Dyer,C.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-JUN-2000) Animal Physiology Research Unit, USDA -
            Agricultural Research Service, University of Missouri, 920 East
            Campus Dr., Room S-143 ASRC, Columbia, MO 65211, USA
FEATURES             Location/Qualifiers
     source          1. .237
                     /organism="Sus scrofa"
                     /mol_type="mRNA"
                     /db_xref="taxon:9823"
                     /tissue_type="brain: hypothalamus"
     CDS             <1. .>237
                     /function="possible role in appetite control and energy
                     regulation"
                     /note="transmembrane protein"
                     /codon_start=2
                     /product="mahogany"
                     /protein_id="AAF82297.1"
                     /db_xref="GI:8980839"
                     /translation="TDCYSCTANTNDCHWCNDHCVPMNHSCTEGQISIAKYENCPKDN
                     PMYYCNKKTSCRSCALDQNCQWEPRNQECIALPE"
ORIGIN

Alignment Scores:
Pred. No.:              1.37e-30        Length:         237
Score:                  42.00           Matches:        42
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            2.94%           Indels:         0
DB:                     4               Gaps:           0

US-09-787-097-12 (1-1429) x AF273610 (1-237)

Qy        744 TyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArg 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        110 TATGAGAATTGCCCAAAGGATAACCCCATGTACTACTGTAACAAGAAGACCAGCTGCAGG 169
```

```
Qy        764 SerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeu 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        170 AGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTG 229

Qy        784 ProGlu 785
              ||||||
Db        230 CCCGAA 235

RESULT 69
AB06968S27
LOCUS       AB06968S27              5036 bp    DNA     linear   ROD 04-APR-2002
DEFINITION  Rattus norvegicus Atrn gene for membrane attractin, exon 29 and
            complete cds.
ACCESSION   AB049248
VERSION     AB049248.1  GI:12275389
KEYWORDS    .
SEGMENT     27 of 27
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1
  AUTHORS   Kuramoto,T., Kitada,K., Inui,T., Sasaki,Y., Ito,K., Hase,T.,
            Kawaguchi,S., Ogawa,Y., Nakao,K., Barsh,G.S., Nagao,M., Ushijima,T.
            and Serikawa,T.
  TITLE     Attractin/mahogany/zitter plays a critical role in myelination of
            the central nervous system
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (2), 559-564 (2001)
  MEDLINE   21143347
   PUBMED   11209055
REFERENCE   2  (bases 1 to 5036)
  AUTHORS   Kuramoto,T. and Ushijima,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-SEP-2000) Takashi Kuramoto, National Cancer Center
            Research Institute, Carcinogenesis Division; Tsukiji 5-1-1,
            Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramot@gan2.ncc.go.jp,
            Tel:81-33542-2511, Fax:81-5565-1753)
FEATURES             Location/Qualifiers
     source          1. .5036
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /chromosome="3"
     gene            join(AB069686.1:3755. .4173,AB069687.1:2348. .2431,
                     AB049224.1:515. .628,AB049225.1:483. .611,
                     AB049226.1:171. .1767,AB049227.1:573. .663,
                     AB049228.1:627. .870,AB049229.1:519. .702,
                     AB049230.1:496. .650,AB049231.1:504. .588,
                     AB049232.1:336. .556,AB049233.1:321. .442,
                     AB049234.1:348. .486,AB049235.1:795. .985,
                     AB049236.1:488. .707,AB049237.1:595. .1516,
                     AB049238.1:445. .582,AB049239.1:107. .200,
                     AB049240.1:422. .474,AB049241.1:117. .211,
                     AB049242.1:507. .585,AB049243.1:366. .523,
                     AB049244.1:508. .569,AB049245.1:522. .600,
                     AB049246.1:500. .607,AB049247.1:548. .662,544. .5036)
                     /gene="Atrn"
     CDS             join(AB069686.1:3755. .4173,AB069687.1:2348. .2431,
                     AB049224.1:515. .628,AB049225.1:483. .611,
                     AB049226.1:171. .376,AB049226.1:1599. .1767,
                     AB049227.1:573. .663,AB049228.1:627. .870,
                     AB049229.1:519. .702,AB049230.1:496. .650,
                     AB049231.1:504. .588,AB049232.1:336. .556,
                     AB049233.1:321. .442,AB049234.1:348. .486,
                     AB049235.1:795. .985,AB049236.1:488. .707,
                     AB049237.1:595. .780,AB049237.1:1283. .1516,
                     AB049238.1:445. .582,AB049239.1:107. .200,
                     AB049240.1:422. .474,AB049241.1:117. .211,
                     AB049242.1:507. .585,AB049243.1:366. .523,
                     AB049244.1:508. .569,AB049245.1:522. .600,
                     AB049246.1:500. .607,AB049247.1:548. .662,544. .668)
                     /gene="Atrn"
                     /codon_start=1
                     /product="membrane attractin"
                     /protein_id="BAB21058.1"
                     /db_xref="GI:12275390"
                     /translation="MVAAAAAAEATEARLRGYTTATAAPAGWKERQHRPCAATGAWRP
                     WPRAGLCLPRVLSRALSPPPLLPLLPLLFSLLLLPLPREAEAAAVAAAVSGSAAAEAK
                     ECDRPCVNGGRCNPGTGQCVCPTGWVGEQCQHCGGRFRLTGSSGFVTDGPGNYKYKTK
                     CTWLIEGQPNKIMRLRFNHFATECSWDHLYVYDGDSIYAPLIAAFSGLIVPERDGNET
                     APEVTVTSGYALLHFFSDAAYNLTGFNITYNFDMCPNNCSGRGECKSSNSSSTVECEC
                     SENWKGESCDIPHCTDNCGFPHRGICNASDTRGCSCFPHWQGPGCSIPVPANQSFWTR
                     EEYSDLKLPRASHKAVVNGNIMWVVGGYMFNHSDYSMVLAYDLASREWLSLNHSVNSV
                     VVRYGHSLALHKDKIYMYGGKIDSTGNVTNELRVFHIHNESWVLLTPKAKDQYAVVGH
                     SAHIVTLSSGRVVMLVIFGHCPLYGYISVVQEYDLEKNTWSILQTQGALVQGGYGHSS
                     VYDHRTKALYVHGGYKAFSANKYRLADDLYRYHVDTQMWTILKDSRFFRYLHTAVIVS
                     GTMLVFGGNTHNDTSMSHGAKCFSSDFMAYDIACDRWSVLPRPELHHDVNRFGHSAVL
                     HNSTMYVFGGFNSLLLSDVLVFTSEQCDAHRSEAACVAAGPGIRCLWDTQSSRCTSWE
                     LATEEQAEKLKSECFSKRTLDHDRCDQHTDCYSCTANTNDCHWCNDHCVPVNHSCTEG
                     QISIAKYDNCPKDNPMYYCNKKTSCRSCALDQNCQWEPRNQECIALPENICGIGWHLV
                     GNSCLKITTAKENYDNAKLSCRNHNAFLASLTSQKKVEFVLKQLRLMQSSQSTSKLTL
                     TPWVGLRKINVSYWCWEDMSPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCIN
                     PLNGSVCERPANHSAKQCRTPCALRTACGECTSSSSECMWCSNMKQCVDSNAYVASFP
                     FGQCMEWYTMSSCPPENCSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKM
                     PSHASTGNVYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCEDLTT
                     GKHCETCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGEECQLCEVE
                     NRYQGNPLKGTCYYTLLIDYQFTFSLSQEDDRYYTAINFVATPDEQNRDLDMFINASK
                     NFNLNITWATSFPAGTQTGEEVPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSN
                     FTWPIKIQIAFSQHSNFMDLVQFFVTFFSCFLSLLLVAAVVWKIKQSCWASRRREQLL
                     REMQQMASRPFASVNVALETDEEPPDLIGGSIKTVPKPIALEPCFGNKAAVLSVFVRL
                     PRGLGGIPPPGQSGLAVASALVDISQQMPIVYKEKSGAVRNRKQQPPAQPGTCI"
     exon            544. .5036
                     /gene="Atrn"
                     /product="membrane attractin"
                     /number=29
ORIGIN

Alignment Scores:
Pred. No.:                1.46e-28        Length:          5036
Score:                    41.00           Matches:         41
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              2.87%           Indels:          0
DB:                       10              Gaps:            0

US-09-787-097-12 (1-1429) x AB06968S27 (1-5036)

Qy       1389 GlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetProIleValTyrLys 1408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 GGTCTTGCTGTGGCCAGTGCCCTGGTGGACATTTCTCAGCAGATGCCAATAGTGTACAAG 602

Qy       1409 GluLysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnProGlyThrCys 1428
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 GAGAAGTCAGGAGCTGTAAGGAACCGGAAGCAGCAGCCCCCTGCACAGCCTGGAACCTGC 662

Qy       1429 Ile 1429
              |||
Db        663 ATT 665

RESULT 70
H009384S27
LOCUS       H009384S27              5790 bp    DNA     linear   PRI 14-DEC-2000
DEFINITION  Homo sapiens attractin precursor (ATRN) gene, exon 30 and complete
            cds, alternatively spliced.
ACCESSION   AF218915
VERSION     AF218915.1  GI:8118081
KEYWORDS    .
SEGMENT     27 of 27
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 5790)
  AUTHORS   Tang,W., Gunn,T.M., McLaughlin,D.F., Barsh,G.S., Schlossman,S.F.
            and Duke-Cohan,J.S.
```

```
  TITLE     Secreted and membrane attractin result from alternative splicing of
            the human ATRN gene
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
  MEDLINE   20283944
   PUBMED   10811918
REFERENCE   2  (bases 1 to 5790)
  AUTHORS   Tang,W., Gunn,T.M., McLaughlin,D.F., Freeman,G.J., Barsh,G.S.,
            Schlossman,S.F. and Duke-Cohan,J.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
            Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
            USA
FEATURES             Location/Qualifiers
     source          1. .5790
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /map="20p13"
     gene            order(AF218889.1:745. .1576,AF218890.1:1. .3711,
                     AF218891.1:1. .3058,AF218892.1:1. .1152,AF218893.1:1. .1232,
                     AF218894.1:1. .1032,AF218895.1:1. .1575,AF218896.1:1. .2018,
                     AF218897.1:1. .3038,AF218898.1:1. .1552,AF218899.1:1. .3589,
                     AF218900.1:1. .1267,AF218901.1:1. .3112,AF218902.1:1. .1728,
                     AF218903.1:1. .3257,AF218904.1:1. .1674,AF218905.1:1. .1675,
                     AF218906.2:1. .966,AF218907.1:1. .1781,AF218908.1:1. .1704,
                     AF218909.1:1. .1976,AF218910.1:1. .1364,AF218911.1:1. .2214,
                     AF218912.1:1. .1556,AF218913.1:1. .1592,AF218914.1:1. .1394,
                     1. .5060)
                     /gene="ATRN"
     mRNA            join(AF218889.1:745. .1233,AF218890.1:2877. .2960,
                     AF218891.1:810. .923,AF218892.1:818. .946,
                     AF218893.1:621. .826,AF218894.1:529. .697,
                     AF218895.1:1343. .1433,AF218896.1:926. .1169,
                     AF218897.1:1114. .1297,AF218897.1:2047. .2201,
                     AF218898.1:659. .743,AF218899.1:1585. .1805,
                     AF218900.1:728. .849,AF218901.1:440. .578,
                     AF218901.1:2163. .2353,AF218902.1:533. .752,
                     AF218903.1:1441. .1626,AF218903.1:2189. .2422,
                     AF218904.1:810. .947,AF218905.1:833. .926,
                     AF218906.2:49. .101,AF218907.1:802. .896,
                     AF218908.1:873. .951,AF218909.1:392. .549,
                     AF218911.1:1348. .1409,AF218912.1:748. .826,
                     AF218913.1:747. .854,AF218914.1:735. .849,660. .5060)
                     /gene="ATRN"
                     /product="membrane attractin precursor"
     CDS             join(AF218889.1:824. .1233,AF218890.1:2877. .2960,
                     AF218891.1:810. .923,AF218892.1:818. .946,
                     AF218893.1:621. .826,AF218894.1:529. .697,
                     AF218895.1:1343. .1433,AF218896.1:926. .1169,
                     AF218897.1:1114. .1297,AF218897.1:2047. .2201,
                     AF218898.1:659. .743,AF218899.1:1585. .1805,
                     AF218900.1:728. .849,AF218901.1:440. .578,
                     AF218901.1:2163. .2353,AF218902.1:533. .752,
                     AF218903.1:1441. .1626,AF218903.1:2189. .2422,
                     AF218904.1:810. .947,AF218905.1:833. .926,
                     AF218906.2:49. .101,AF218907.1:802. .896,
                     AF218908.1:873. .951,AF218909.1:392. .549,
                     AF218911.1:1348. .1409,AF218912.1:748. .826,
                     AF218913.1:747. .854,AF218914.1:735. .849,660. .784)
                     /gene="ATRN"
                     /note="Method: conceptual translation with partial peptide
                     sequencing"
                     /codon_start=1
                     /product="membrane attractin precursor"
                     /protein_id="AAF72881.1"
                     /db_xref="GI:8118082"
                     /translation="MVAAAAATEARLRRRTAATAALAGRSGGPHWDWDVTRAGRPGLG
                     AGLRLPRLLSPPLRPRLLLLLLLLSPPLLLLLPCEAEAAAAAAAVSGSAAAEAKECD
                     RPCVNGGRCNPGTGQCVCPAGWVGEQCQHCGGRFRLTGSSGFVTDGPGNYKYKTKCTW
                     LIEGQPNRIMRLRFNHFATECSWDHLYVYDGDSIYAPLVAAFSGLIVPERDGNETVPE
                     VVATSGYALLHFFSDAAYNLTGFNITYSFDMCPNNCSGRGECKISNSSDTVECECSEN
                     WKGEACDIPHCTDNCGFPHRGICNSSDVRGCSCFSDWQGPGCSVPVPANQSFWTREEY
                     SNLKLPRASHKAVVNGNIMWVVGGYMFNHSDYNMVLAYDLASREWLPLNRSVNNVVVR
                     YGHSLALYKDKIYMYGGKIDSTGNVTNELRVFHIHNESWVLLTPKAKEQYAVVGHSAH
                     IVTLKNGRVVMLVIFGHCPLYGYISNVQEYDLDKNTWSILHTQGALVQGGYGHSSVYD
                     HRTRALYVHGGYKAFSANKYRLADDLYRYDVDTQMWTILKDSRFFRYLHTAVIVSGTM
                     LVFGGNTHNDTSMSHGAKCFSSDFMAYDIACDRWSVLPRPDLHHDVNRFGHSAVLHNS
                     TMYVFGGFNSLLLSDILVFTSEQCDAHRSEAACLAAGPGIRCVWNTGSSQCISWALAT
                     DEQEEKLKSECFSKRTLDHDRCDQHTDCYSCTANTNDCHWCNDHCVPRNHSCSEGQIS
                     IFRYENCPKDNPMYYCNKKTSCRSCALDQNCQWEPRNQECIALPENICGIGWHLVGNS
                     CLKITTAKENYDNAKLFCRNHNALLASLTTQKKVEFVLKQLRIMQSSQSMSKLTLTPW
                     VGLRKINVSYWCWEDMSPFTNSLLQWMPSEPSDAGFCGILSEPSTRGLKAATCINPLN
                     GSVCERPANHSAKQCRTPCALRTACGDCTSGSSECMWCSNMKQCVDSNAYVASFPFGQ
                     CMEWYTMSTCPPENCSGYCTCSHCLEQPGCGWCTDPSNTGKGKCIEGSYKGPVKMPSQ
                     APTGNFYPQPLLNSSMCLEDSRYNWSFIHCPACQCNGHSKCINQSICEKCENLTTGKH
                     CETCISGFYGDPTNGGKCQPCKCNGHASLCNTNTGKCFCTTKGVKGDECQLCEVENRY
                     QGNPLRGTCYYTLLIDYQFTFSLSQKDDRYYTAINFVATPDEQNRDLDMFINASKNFN
                     LNITWAASFSAGTQAGEEMPVVSKTNIKEYKDSFSNEKFDFRNHPNITFFVYVSNFTW
                     PIKIQIAFSQHSNFMDLVQFFVTFFSCFLSLLLVAAVVWKIKQSCWASRRREQLLREM
                     QQMASRPFASVNVALETDEEPPDLIGGSIKTVPKPIALEPCFGNKAAVLSVFVRLPRG
                     LGGIPPPGQSGLAVASALVDISQQMPIVYKEKSGAVRNRKQQPPAQPGTCI"
     mat_peptide     join(AF218889.1:1073. .1233,AF218890.1:2877. .2960,
                     AF218891.1:810. .923,AF218892.1:818. .946,
                     AF218893.1:621. .826,AF218894.1:529. .697,
                     AF218895.1:1343. .1433,AF218896.1:926. .1169,
                     AF218897.1:1114. .1297,AF218897.1:2047. .2201,
                     AF218898.1:659. .743,AF218899.1:1585. .1805,
                     AF218900.1:728. .849,AF218901.1:440. .578,
                     AF218901.1:2163. .2353,AF218902.1:533. .752,
                     AF218903.1:1441. .1626,AF218903.1:2189. .2422,
                     AF218904.1:810. .947,AF218905.1:833. .926,
                     AF218906.2:49. .101,AF218907.1:802. .896,
                     AF218908.1:873. .951,AF218909.1:392. .549,
                     AF218911.1:1348. .1409,AF218912.1:748. .826,
                     AF218913.1:747. .854,AF218914.1:735. .849,660. .781)
                     /gene="ATRN"
                     /product="membrane attractin"
     exon            660. .5060
                     /gene="ATRN"
                     /number=30
     polyA_signal    5038. .5043
                     /gene="ATRN"
ORIGIN

Alignment Scores:
Pred. No.:                1.63e-28        Length:          5790
Score:                    41.00           Matches:         41
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              2.87%           Indels:          0
DB:                       9               Gaps:            0

US-09-787-097-12 (1-1429) x H009384S27 (1-5790)

Qy       1389 GlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetProIleValTyrLys 1408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        659 GGTCTTGCTGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGATAGTGTACAAG 718

Qy       1409 GluLysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnProGlyThrCys 1428
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 GAGAAGTCAGGAGCCGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGCCTGGGACCTGC 778

Qy       1429 Ile 1429
              |||
Db        779 ATC 781

RESULT 71
MMMAHNY2
LOCUS       MMMAHNY2                7437 bp    DNA     linear   ROD 31-MAR-1999
DEFINITION  Mus musculus mahogany (mg) gene, exons and partial cds.
ACCESSION   AF120318
VERSION     AF120318.1  GI:4545288
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Mus musculus (house mouse)
```

```
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 7437)
  AUTHORS   Gunn,T.M., Miller,K.A., He,L., Hyman,R.W., Davis,R.W., Azarani,A.,
            Schlossman,S.F., Duke-Cohan,J.S. and Barsh,G.S.
  TITLE     The mouse mahogany locus encodes a transmembrane form of human
            attractin
  JOURNAL   Nature 398 (6723), 152-156 (1999)
  MEDLINE   99184160
   PUBMED   10086356
REFERENCE   2  (bases 1 to 7437)
  AUTHORS   Gunn,T.M., Miller,K.A., He,L., Hyman,R.W., Davis,R.W., Azarani,A.,
            Schlossman,S.F., Duke-Cohan,J.S. and Barsh,G.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JAN-1999) HHMI, Stanford, Beckman Center B271,
            Stanford, CA 94305-5323, USA
FEATURES             Location/Qualifiers
     source          1. .7437
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="2"
                     /clone="BAC389B9"
     gene            order(AF120317.1:<1. .9085,1. .>7437)
                     /gene="mg"
     mRNA            join(AF120317.1:<1870. .1931,AF120317.1:6879. .6957,
                     282. .389,2342. .2456,5156. .>7437)
                     /gene="mg"
                     /product="mahogany"
     CDS             join(AF120317.1:<1870. .1931,AF120317.1:6879. .6957,
                     282. .389,2342. .2456,5156. .5280)
                     /gene="mg"
                     /note="corresponding cDNA sequence found in GenBank
                     Accession Number AF119821"
                     /codon_start=1
                     /product="mahogany"
                     /protein_id="AAD22476.1"
                     /db_xref="GI:4545290"
                     /translation="IAFSQHSNFMDLVQFFVTFFSCFLSLLLVAAVVWKIKQSCWASR
                     RREQLLREMQQMASRPFASVNVALETDEEPPDLIGGSIKTVPKPIALEPCFGNKAAVL
                     SVFVRLPRGLGGIPPPGQSGLAVASALVDISQQMPIVYKEKSGAVRNRKQQPPAQPGT
                     CI"
     exon            282. .389
                     /gene="mg"
     exon            2342. .2456
                     /gene="mg"
     exon            5156. .>7437
                     /gene="mg"
     3'UTR           5281. .>7437
                     /gene="mg"
ORIGIN

Alignment Scores:
Pred. No.:              1.99e-28        Length:         7437
Score:                  41.00           Matches:        41
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            2.87%           Indels:         0
DB:                     10              Gaps:           0

US-09-787-097-12 (1-1429) x MMMAHNY2 (1-7437)

Qy       1389 GlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetProIleValTyrLys 1408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5155 GGTCTCGCTGTGGCCAGTGCCCTGGTGGACATTTCTCAGCAGATGCCAATAGTGTACAAG 5214

Qy       1409 GluLysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnProGlyThrCys 1428
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5215 GAGAAGTCAGGAGCTGTAAGAAACCGGAAGCAGCAGCCGCCTGCACAGCCTGGAACCTGC 5274

Qy       1429 Ile 1429
              |||
Db       5275 ATT 5277

RESULT 72
AC017113/c
LOCUS       AC017113               206329 bp    DNA     linear   HTG 07-JUL-2000
DEFINITION  Homo sapiens chromosome 20 clone RP11-574H7, WORKING DRAFT
            SEQUENCE, 31 unordered pieces.
ACCESSION   AC017113
VERSION     AC017113.4  GI:8570339
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 206329)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 206329)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-DEC-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Jun 17, 2000 this sequence version replaced gi:6855220.

            ----------------- Genome Center -----------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            -------------- Project Information --------------
            Center project name: H_NH0574H07
            -------------- Summary Statistics  --------------
            Sequencing vector: M13; 91%
            Sequencing vector: plasmid; 9%
            Chemistry: Dye-primer ET; 91% of reads
            Chemistry: Dye-terminator Big Dye; 9% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 184921 bases at least Q40
            Consensus quality: 192397 bases at least Q30
            Consensus quality: 195862 bases at least Q20
            Insert size: 237000; agarose-fp
            Insert size: 203329; sum-of-contigs
            Quality coverage: 3.84 in Q20 bases; agarose-fp
            Quality coverage: 3.66 in Q20 bases; sum-of-contigs
            ------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 31 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1539: contig of 1539 bp in length
            *     1540     1639: gap of unknown length
            *     1640     3259: contig of 1620 bp in length
            *     3260     3359: gap of unknown length
            *     3360     5400: contig of 2041 bp in length
            *     5401     5500: gap of unknown length
            *     5501     7783: contig of 2283 bp in length
            *     7784     7883: gap of unknown length
            *     7884    11241: contig of 3358 bp in length
            *    11242    11341: gap of unknown length
            *    11342    13482: contig of 2141 bp in length
            *    13483    13582: gap of unknown length
            *    13583    16915: contig of 3333 bp in length
            *    16916    17015: gap of unknown length
            *    17016    21346: contig of 4331 bp in length
            *    21347    21446: gap of unknown length
            *    21447    25103: contig of 3657 bp in length
            *    25104    25203: gap of unknown length
            *    25204    28415: contig of 3212 bp in length
```

```
            *    28416    28515: gap of unknown length
            *    28516    32873: contig of 4358 bp in length
            *    32874    32973: gap of unknown length
            *    32974    36496: contig of 3523 bp in length
            *    36497    36596: gap of unknown length
            *    36597    40649: contig of 4053 bp in length
            *    40650    40749: gap of unknown length
            *    40750    44878: contig of 4129 bp in length
            *    44879    44978: gap of unknown length
            *    44979    50353: contig of 5375 bp in length
            *    50354    50453: gap of unknown length
            *    50454    55946: contig of 5493 bp in length
            *    55947    56046: gap of unknown length
            *    56047    61420: contig of 5374 bp in length
            *    61421    61520: gap of unknown length
            *    61521    67819: contig of 6299 bp in length
            *    67820    67919: gap of unknown length
            *    67920    72327: contig of 4408 bp in length
            *    72328    72427: gap of unknown length
            *    72428    77293: contig of 4866 bp in length
            *    77294    77393: gap of unknown length
            *    77394    82626: contig of 5233 bp in length
            *    82627    82726: gap of unknown length
            *    82727    90186: contig of 7460 bp in length
            *    90187    90286: gap of unknown length
            *    90287    96584: contig of 6298 bp in length
            *    96585    96684: gap of unknown length
            *    96685   106019: contig of 9335 bp in length
            *   106020   106119: gap of unknown length
            *   106120   113530: contig of 7411 bp in length
            *   113531   113630: gap of unknown length
            *   113631   125151: contig of 11521 bp in length
            *   125152   125251: gap of unknown length
            *   125252   137677: contig of 12426 bp in length
            *   137678   137777: gap of unknown length
            *   137778   151263: contig of 13486 bp in length
            *   151264   151363: gap of unknown length
            *   151364   165370: contig of 14007 bp in length
            *   165371   165470: gap of unknown length
            *   165471   184704: contig of 19234 bp in length
            *   184705   184804: gap of unknown length
            *   184805   206329: contig of 21525 bp in length.
FEATURES             Location/Qualifiers
     source          1. .206329
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /clone="RP11-574H7"
ORIGIN

Alignment Scores:
Pred. No.:              2.86e-27        Length:         206329
Score:                  41.00           Matches:        41
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            2.87%           Indels:         0
DB:                     2               Gaps:           0

US-09-787-097-12 (1-1429) x AC017113 (1-206329)

Qy       1389 GlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetProIleValTyrLys 1408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     159327 GGTCTTGCTGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGATAGTGTACAAG 159268

Qy       1409 GluLysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnProGlyThrCys 1428
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     159267 GAGAAGTCAGGAGCCGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGCCTGGGACCTGC 159208

Qy       1429 Ile 1429
              |||
Db     159207 ATC 159205
```

```
RESULT 73
AF415225/c
LOCUS       AF415225               223026 bp    DNA     linear   HTG 11-JUL-2002
DEFINITION  Mus musculus chromosome 2 clone CITB BAC 389B9 strain 129, ***
            SEQUENCING IN PROGRESS ***, 25 unordered pieces.
ACCESSION   AF415225
VERSION     AF415225.1  GI:15741219
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 223026)
  AUTHORS   Gunn,T.M., Azarani,A., Kim,P.H., Hyman,R.W., Davis,R.W. and
            Barsh,G.S.
  TITLE     Identification and preliminary characterization of mouse Adam33
  JOURNAL   BMC Genet. 3 (1), 2 (2002)
   PUBMED   11897009
REFERENCE   2  (bases 1 to 223026)
  AUTHORS   Gunn,T.M., Azarani,A., Kim,P., Hyman,R.W., Davis,R.W. and
            Barsh,G.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-AUG-2001) HHMI, Stanford University, Beckman Center,
            Stanford University Medical Center, Stanford, CA 94305, USA
COMMENT     Sequenced to 6X coverage. Clone contains genes: Atrn, Adam33, Sn,
            Cdc25b, Cenpb, Gfra4 and EST clusters represented by IMAGE clones
            3868736, 388220, 775311, 614556, 427645, 807697 and 514017.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 25 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1      764: contig of 764 bp in length
            *     765      864: gap of unknown length
            *     865     1731: contig of 867 bp in length
            *    1732     1831: gap of unknown length
            *    1832     3030: contig of 1199 bp in length
            *    3031     3130: gap of unknown length
            *    3131     4476: contig of 1346 bp in length
            *    4477     4576: gap of unknown length
            *    4577     5597: contig of 1021 bp in length
            *    5598     5697: gap of unknown length
            *    5698     7086: contig of 1389 bp in length
            *    7087     7186: gap of unknown length
            *    7187     8614: contig of 1428 bp in length
            *    8615     8714: gap of unknown length
            *    8715    10700: contig of 1986 bp in length
            *   10701    10800: gap of unknown length
            *   10801    13233: contig of 2433 bp in length
            *   13234    13333: gap of unknown length
            *   13334    16153: contig of 2820 bp in length
            *   16154    16253: gap of unknown length
            *   16254    20801: contig of 4548 bp in length
            *   20802    20901: gap of unknown length
            *   20902    23601: contig of 2700 bp in length
            *   23602    23701: gap of unknown length
            *   23702    27816: contig of 4115 bp in length
            *   27817    27916: gap of unknown length
            *   27917    37125: contig of 9209 bp in length
            *   37126    37225: gap of unknown length
            *   37226    44951: contig of 7726 bp in length
            *   44952    45051: gap of unknown length
            *   45052    54136: contig of 9085 bp in length
            *   54137    54236: gap of unknown length
            *   54237    61673: contig of 7437 bp in length
            *   61674    61773: gap of unknown length
            *   61774    71077: contig of 9304 bp in length
            *   71078    71177: gap of unknown length
            *   71178    81674: contig of 10497 bp in length
            *   81675    81774: gap of unknown length
```

```
          *    81775    92951: contig of 11177 bp in length
          *    92952    93051: gap of unknown length
          *    93052   107739: contig of 14688 bp in length
          *   107740   107839: gap of unknown length
          *   107840   126256: contig of 18417 bp in length
          *   126257   126356: gap of unknown length
          *   126357   145400: contig of 19044 bp in length
          *   145401   145500: gap of unknown length
          *   145501   175039: contig of 29539 bp in length
          *   175040   175139: gap of unknown length
          *   175140   223026: contig of 47887 bp in length.
FEATURES             Location/Qualifiers
     source          1. .223026
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="129"
                     /db_xref="taxon:10090"
                     /chromosome="2"
                     /clone="CITB BAC 389B9"
ORIGIN

Alignment Scores:
Pred. No.:              3.04e-27        Length:         223026
Score:                  41.00           Matches:        41
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            2.87%           Indels:         0
DB:                     2               Gaps:           0

US-09-787-097-12 (1-1429) x AF415225 (1-223026)

Qy       1389 GlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetProIleValTyrLys 1408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      56519 GGTCTCGCTGTGGCCAGTGCCCTGGTGGACATTTCTCAGCAGATGCCAATAGTGTACAAG 56460

Qy       1409 GluLysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnProGlyThrCys 1428
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      56459 GAGAAGTCAGGAGCTGTAAGAAACCGGAAGCAGCAGCCGCCTGCACAGCCTGGAACCTGC 56400

Qy       1429 Ile 1429
              |||
Db      56399 ATT 56397

RESULT 74
H009384S12
LOCUS       H009384S12              1267 bp    DNA     linear   PRI 14-DEC-2000
DEFINITION  Homo sapiens attractin precursor (ATRN) gene, exon 13.
ACCESSION   AF218900
VERSION     AF218900.1  GI:8118066
KEYWORDS    .
SEGMENT     12 of 27
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1267)
  AUTHORS   Tang,W., Gunn,T.M., McLaughlin,D.F., Barsh,G.S., Schlossman,S.F.
            and Duke-Cohan,J.S.
  TITLE     Secreted and membrane attractin result from alternative splicing of
            the human ATRN gene
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
  MEDLINE   20283944
   PUBMED   10811918
REFERENCE   2  (bases 1 to 1267)
  AUTHORS   Tang,W., Gunn,T.M., McLaughlin,D.F., Freeman,G.J., Barsh,G.S.,
            Schlossman,S.F. and Duke-Cohan,J.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
            Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
            USA
FEATURES             Location/Qualifiers
     source          1. .1267
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /map="20p13"
     exon            728. .849
                     /gene="ATRN"
                     /number=13
ORIGIN

Alignment Scores:
Pred. No.:              4.43e-28        Length:         1267
Score:                  40.00           Matches:        40
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            2.80%           Indels:         0
DB:                     9               Gaps:           0

US-09-787-097-12 (1-1429) x H009384S12 (1-1267)

Qy        699 LeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsn 718
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        730 CTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAAT 789

Qy        719 AspCysHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGluGlyGln 738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        790 GACTGCCACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAG 849

RESULT 75
AB06968S13
LOCUS       AB06968S13              1071 bp    DNA     linear   ROD 04-APR-2002
DEFINITION  Rattus norvegicus Atrn gene for membrane attractin, exon 14.
ACCESSION   AB049234
VERSION     AB049234.1  GI:12275375
KEYWORDS    .
SEGMENT     13 of 27
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1
  AUTHORS   Kuramoto,T., Kitada,K., Inui,T., Sasaki,Y., Ito,K., Hase,T.,
            Kawaguchi,S., Ogawa,Y., Nakao,K., Barsh,G.S., Nagao,M., Ushijima,T.
            and Serikawa,T.
  TITLE     Attractin/mahogany/zitter plays a critical role in myelination of
            the central nervous system
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (2), 559-564 (2001)
  MEDLINE   21143347
   PUBMED   11209055
REFERENCE   2  (bases 1 to 1071)
  AUTHORS   Kuramoto,T. and Ushijima,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-SEP-2000) Takashi Kuramoto, National Cancer Center
            Research Institute, Carcinogenesis Division; Tsukiji 5-1-1,
            Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramot@gan2.ncc.go.jp,
            Tel:81-33542-2511, Fax:81-5565-1753)
FEATURES             Location/Qualifiers
     source          1. .1071
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /chromosome="3"
     exon            348. .486
                     /gene="Atrn"
                     /product="membrane attractin"
                     /note="CDS is reported in Acc# AB049248"
                     /number=14
ORIGIN

Alignment Scores:
Pred. No.:              3.56e-27        Length:         1071
Score:                  39.00           Matches:        39
Percent Similarity:     100.00%         Conservative:   0
```

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.73% Indels: 0
DB: 10 Gaps: 0

US-09-787-097-12 (1-1429) x AB06968S13 (1-1071)

```
Qy        746 AsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArgSerCys 765
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 AATTGCCCCAAGGATAACCCCATGTACTACTGCAATAAGAAAACCAGCTGCAGGAGCTGT 428

Qy        766 AlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuPro 784
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        429 GCCCTAGATCAGAATTGCCAGTGGGAACCTCGAAATCAAGAGTGCATCGCCCTGCCC 485
```

Search completed: March 1, 2004, 12:04:58
Job time : 15574 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 1, 2004, 02:00:44 ; Search time 1075 Seconds
(without alignments)
5647.140 Million cell updates/sec

Title: US-09-787-097-12
Perfect score: 1429
Sequence: 1 MVAAAAATEARLRRRTAATA.....KSGAVRNRKQQPPAQPGTCI 1429

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Word size: 1

Total number of hits satisfying chosen parameters: 6745146

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09787097/runat_26022004_063014_23737/app_query.fasta_1.1607
-DB=N_Geneseq_29Jan04 -QFMT=fastap -SUFFIX=oligo.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=100 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=75 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09787097_@CGN_1_1_1137_@runat_26022004_063014_23737 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1429	100.0	4290	3	AAZ52276	Aaz52276 Human mem
2	1429	100.0	8561	5	AAS72659	Aas72659 DNA encod
3	1267	88.7	3819	3	AAZ52277	Aaz52277 Human sol
4	1141	79.8	8590	3	AAZ91920	Aaz91920 Human mah
5	1123	78.6	4068	3	AAZ52275	Aaz52275 Human mem
6	979	68.5	4072	3	AAZ91921	Aaz91921 Human mah
7	874	61.2	3597	3	AAZ52274	Aaz52274 Human sol
8	874	61.2	3597	5	AAS72657	Aas72657 DNA encod
9	874	61.2	3597	9	ADD70997	Add70997 Human att
10	496	34.7	2625	3	AAZ91922	Aaz91922 Human mah
11	467	32.7	6373	3	AAZ91919	Aaz91919 Human mah
12	289	20.2	3490	5	AAS72660	Aas72660 DNA encod
13	206	14.4	8739	9	ADB53249	Adb53249 Primary r
14	181	12.7	8827	3	AAZ91916	Aaz91916 Wild type
15	142	9.9	3165	5	AAS78781	Aas78781 DNA encod
16	142	9.9	3313	5	AAS76294	Aas76294 DNA encod
17	110	7.7	671	5	AAS72658	Aas72658 DNA encod
18	104	7.3	668	4	AAS02409	Aas02409 Human sec
19	94	6.6	2419	3	AAZ91917	Aaz91917 Murine ma
c 20	92	6.4	481	4	AAL14686	Aal14686 Human bre
21	92	6.4	1012	2	AAX84568	Aax84568 Biorhythm
c 22	82	5.7	246	6	ABK38928	Abk38928 cDNA enco
c 23	82	5.7	246	7	ACA11257	Aca11257 Human lun
c 24	82	5.7	246	7	ACA02443	Aca02443 Lung canc
c 25	79	5.5	482	4	AAL23548	Aal23548 Human bre
26	77	5.4	207433	5	ABZ72040	Abz72040 Gene 216
27	77	5.4	207433	7	ABX74891	Abx74891 BAC1098L2
28	76	5.3	1051	3	AAZ91918	Aaz91918 Murine ma
29	73	5.1	90050	3	AAZ91925	Aaz91925 Wild type
30	69	4.8	531	4	AAI43425	Aai43425 Probe #12
31	69	4.8	531	6	ABS11523	Abs11523 Human gen
32	64	4.5	193	4	AAI56405	Aai56405 Probe #25
33	64	4.5	193	6	ABS23881	Abs23881 Human gen
34	64	4.5	292	3	AAA45505	Aaa45505 Human sec
c 35	60	4.2	182	4	AAI24099	Aai24099 Probe #14
c 36	60	4.2	182	4	ABA69223	Aba69223 Human foe
c 37	60	4.2	182	4	AAI49392	Aai49392 Probe #18
c 38	60	4.2	182	4	ABA51213	Aba51213 Human bre
c 39	60	4.2	182	4	ABA36140	Aba36140 Probe #14
c 40	60	4.2	182	4	AAK43324	Aak43324 Human bon
c 41	60	4.2	182	4	AAK17517	Aak17517 Human bra
c 42	60	4.2	182	4	ABS42949	Abs42949 Human liv
c 43	60	4.2	182	5	AAI09674	Aai09674 Probe #96
c 44	60	4.2	182	6	ABS17404	Abs17404 Human gen
c 45	57	4.0	375	4	AAI14910	Aai14910 Probe #48
c 46	57	4.0	375	4	ABA56648	Aba56648 Human foe
c 47	57	4.0	375	4	AAI36262	Aai36262 Probe #49
c 48	57	4.0	375	4	ABA46100	Aba46100 Human bre
c 49	57	4.0	375	4	ABA26254	Aba26254 Probe #47
c 50	57	4.0	375	4	AAK30307	Aak30307 Human bon
c 51	57	4.0	375	4	AAK04785	Aak04785 Human bra
c 52	57	4.0	375	4	ABS29950	Abs29950 Human liv
c 53	57	4.0	375	5	AAI04689	Aai04689 Probe #46
c 54	57	4.0	375	6	ABS04892	Abs04892 Human gen
c 55	51	3.6	226	3	AAZ91926	Aaz91926 Wild type
56	49	3.4	2267	5	AAS69062	Aas69062 DNA encod
c 57	46	3.2	402	4	AAI13789	Aai13789 Probe #37
c 58	46	3.2	402	4	ABA55499	Aba55499 Human foe
c 59	46	3.2	402	4	AAI35149	Aai35149 Probe #38
c 60	46	3.2	402	4	ABA45027	Aba45027 Human bre
c 61	46	3.2	402	4	ABA25215	Aba25215 Probe #36
c 62	46	3.2	402	4	AAK29197	Aak29197 Human bon
c 63	46	3.2	402	4	AAK03733	Aak03733 Human bra
c 64	46	3.2	402	4	ABS28814	Abs28814 Human liv
c 65	46	3.2	402	5	AAI03665	Aai03665 Probe #36
c 66	46	3.2	402	6	ABS03749	Abs03749 Human gen
c 67	45	3.1	137	4	AAI23009	Aai23009 Probe #12
c 68	45	3.1	137	4	ABA68099	Aba68099 Human foe
c 69	45	3.1	137	4	AAI48309	Aai48309 Probe #16
c 70	45	3.1	137	4	ABA50165	Aba50165 Human bre
c 71	45	3.1	137	4	ABA35129	Aba35129 Probe #13
c 72	45	3.1	137	4	AAK42239	Aak42239 Human bon
c 73	45	3.1	137	4	AAK16486	Aak16486 Human bra
c 74	45	3.1	137	4	ABS41845	Abs41845 Human liv
c 75	45	3.1	137	5	AAI08671	Aai08671 Probe #86
c 76	45	3.1	137	6	ABS16289	Abs16289 Human gen
77	33	2.3	5973	3	AAZ91924	Aaz91924 Wild type
c 78	30	2.1	16956	3	AAZ91923	Aaz91923 Wild type
79	29	2.0	4134	7	ABZ58710	Abz58710 Murine ho
80	29	2.0	4140	7	ABZ58709	Abz58709 Human hom
81	29	2.0	6517	4	AAH98748	Aah98748 Human EST

	82	22	1.5	657	7	ABZ58714	Abz58714 Upregulat
c	83	22	1.5	845	6	ABQ52396	Abq52396 Oligonucl
	84	22	1.5	845	6	ABQ52397	Abq52397 Oligonucl
	85	22	1.5	982	6	ABL39784	Abl39784 Human NS
	86	22	1.5	1654	7	ABX34730	Abx34730 Human mdd
	87	22	1.5	3609	6	ABS63434	Abs63434 Human cDN
c	88	22	1.5	5312	6	ABL32197	Abl32197 Human imm
	89	20	1.4	60	6	ABN42673	Abn42673 Human spl
	90	20	1.4	503	5	AAS77314	Aas77314 DNA encod
	91	19	1.3	6532	5	AAS86766	Aas86766 DNA encod
	92	18	1.3	65	6	ABN57242	Abn57242 Mouse spl
c	93	17	1.2	51	4	AAL33285	Aal33285 Human SNP
	94	16	1.1	233	6	ABQ97182	Abq97182 Mouse ES
c	95	16	1.1	423	5	AAS90685	Aas90685 DNA encod
c	96	16	1.1	1104	5	AAS90708	Aas90708 DNA encod
c	97	16	1.1	1788	5	AAS70516	Aas70516 DNA encod
	98	15	1.0	1383	6	AAL54064	Aal54064 cDNA of h
c	99	15	1.0	4461	4	ABL14674	Abl14674 Drosophil
c	100	14	1.0	510	5	AAS90687	Aas90687 DNA encod

ALIGNMENTS

```
RESULT 1
AAZ52276
ID   AAZ52276 standard; cDNA; 4290 BP.
XX
AC   AAZ52276;
XX
DT   18-JUL-2000  (first entry)
XX
DE   Human membrane attractin-2 cDNA.
XX
KW   Human; membrane attractin-2; immune response; macrophage; monocyte;
KW   T cell; immunostimulant; immunosuppressed patient; cancer;
KW   immunodeficiency syndrome; transplant; autoimmune disease; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .4290
FT                   /*tag=  a
FT                   /product= "Membrane attractin-2"
XX
PN   WO200015651-A1.
XX
PD   23-MAR-2000.
XX
PF   14-SEP-1999;   99WO-US020948.
XX
PR   14-SEP-1998;   98US-0100137P.
XX
PA   (DAND ) DANA FARBER CANCER INST INC.
XX
PI   Duke-Cohan JS,  Schlossman SF;
XX
DR   WPI; 2000-271373/23.
DR   P-PSDB; AAY70691.
XX
PT   Isolated nucleic acids encoding human attractin polypeptides useful for
PT   enhancing immune responses.
XX
PS   Claim 3; Fig 14; 120pp; English.
XX
CC   The patent discloses four forms of human attractin polypeptides which
CC   enhance immune response by promoting macrophage and monocyte spreading in
CC   the presence of T cells. These include soluble attractin-1 and -2 and
CC   membrane attractin-1 and -2. These various forms of attractin are encoded
CC   by alternatively spliced mRNA molecule transcribed from a single gene.
CC   The present sequence is a cDNA encoding human membrane attractin-2.
CC   Membrane attractin differs from soluble attractin in having a
CC   transmembrane domain and a cytoplasmic domain at the C-terminal.
CC   Attractin-2 differs from attractin-1 in having a 74 amino acid insert in
CC   the N-terminal. This sequence can be used to enhance immune response in
CC   immunosuppressed patients such as those undergoing chemo- and radio-
CC   therapy treatment for cancer or those suffering from common variable
CC   immunodeficiency syndrome. The proteins may also be used to screen
CC   modulators (agonists and antagonists) of immune response which may also
CC   be used to regulate immune reactions. Attractin antibodies can be used to
CC   inhibit immune response in transplant recipients or patients afflicted
CC   with autoimmune disease
XX
SQ   Sequence 4290 BP; 1060 A; 1025 C; 1120 G; 1085 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                0          Length:          4290
Score:                    1429.00    Matches:         1429
Percent Similarity:       100.00%    Conservative:    0
Best Local Similarity:    100.00%    Mismatches:      0
Query Match:              100.00%    Indels:          0
DB:                       3          Gaps:            0

US-09-787-097-12 (1-1429) x AAZ52276 (1-4290)

Qy          1 MetValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 60

Qy         21 AlaLeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArg 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGGGACTGGGACGTGACCAGGGCTGGGAGG 120

Qy         41 ProGlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSerProProLeuArgProArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGGGGCTGGGGGCCGGGCTGCGCCTCCCGCGGCTGCTGTCTCCACCGCTGCGGCCACGG 180

Qy         61 LeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeuProCysGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGCTGCTGCTGCTGTTGTTGCTCCCGCCGCCGCTGTTGCTGCTGCTGCTGCCCTGTGAG 240

Qy         81 AlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCGAGGCCGCGGCGGCGGCGGCGGCGGTGTCGGGCTCAGCCGCAGCCGAGGCCAAGGAA 300

Qy        101 CysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCys 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGTGACCGGCCCTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGC 360

Qy        121 ProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGly 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGA 420

Qy        141 SerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrp 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGG 480

Qy        161 LeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAG 540

Qy        181 CysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAla 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCT 600

Qy        201 AlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGluValVal 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTT 660

Qy        221 AlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGly 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGA 720

Qy        241 PheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCys 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGT 780
```


Qy 261 LysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsnTrpLysGlyGlu 280
Db 781 AAGATCAGTAATAGCAGCGAAACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAA 840
Qy 281 AlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsn 300
Db 841 GCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAAT 900
Qy 301 SerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerVal 320
Db 901 TCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTT 960
Qy 321 ProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuPro 340
Db 961 CCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCC 1020
Qy 341 ArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMet 360
Db 1021 AGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATG 1080
Qy 361 PheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeu 380
Db 1081 TTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTT 1140
Qy 381 ProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyr 400
Db 1141 CCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATAC 1200
Qy 401 LysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsnValThrAsnGlu 420
Db 1201 AAGGATAAAATTTACATGTATGGAGGAAAAATTGATCCAACTGGGAATGTGACCAATGAG 1260
Qy 421 LeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGlu 440
Db 1261 TTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGGAG 1320
Qy 441 GlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValVal 460
Db 1321 CAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTC 1380
Qy 461 MetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyr 480
Db 1381 ATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATAT 1440
Qy 481 AspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGly 500
Db 1441 GATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGT 1500
Qy 501 TyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyr 520
Db 1501 TACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTAC 1560
Qy 521 LysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAsp 540
Db 1561 AAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGAT 1620
Qy 541 ThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaVal 560
Db 1621 ACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTG 1680
Qy 561 IleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSer 580
Db 1681 ATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGC 1740
Qy 581 HisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrp 600
Db 1741 CATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGG 1800
Qy 601 SerValLeuProArgProAspLeuHisHisAspValAsnArgPheGlyHisSerAlaVal 620
Db 1801 TCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTC 1860
Qy 621 LeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIle 640
Db 1861 TTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATC 1920
Qy 641 LeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAla 660
Db 1921 CTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCA 1980
Qy 661 GlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeu 680
Db 1981 GGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTG 2040
Qy 681 AlaThrAspGluGlnGluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAsp 700
Db 2041 GCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGAC 2100
Qy 701 HisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCys 720
Db 2101 CATGACAGATGTGACCAGCACACAGATTGTTACAGCTGTACAGCCAACACCAATGACTGC 2160
Qy 721 HisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSer 740
Db 2161 CACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCC 2220
Qy 741 IlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThr 760
Db 2221 ATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCCATGTACTACTGTAACAAGAAGACC 2280
Qy 761 SerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCys 780
Db 2281 AGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGC 2340
Qy 781 IleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeu 800
Db 2341 ATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTG 2400
Qy 801 LysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsn 820
Db 2401 AAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAAT 2460
Qy 821 AlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArg 840
Db 2461 GCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGA 2520
Qy 841 IleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArg 860
Db 2521 ATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGG 2580
Qy 861 LysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeu 880
Db 2581 AAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTA 2640
Qy 881 GlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSer 900
Db 2641 CAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGT 2700
Qy 901 ThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArg 920
Db 2701 ACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGG 2760
Qy 921 ProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGly 940
Db 2761 CCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGA 2820
Qy 941 AspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAsp 960
Db 2821 GATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGAC 2880
Qy 961 SerAsnAlaTyrValAlaSerProProGlyGlnCysMetGluTrpTyrThrMetSer 980
Db 2881 TCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGC 2940
Qy 981 ThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPro 1000

```
Db       2941 ACCTGCCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCA 3000

Qy       1001 GlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySer 1020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 GGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCC 3060

Qy       1021 TyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnPro 1040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 TATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCC 3120

Qy       1041 LeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCys 1060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 CTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGT 3180

Qy       1061 ProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCys 1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 CCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGT 3240

Qy       1081 GluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspPro 1100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 GAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCC 3300

Qy       1101 ThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThr 1120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 ACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACC 3360

Qy       1121 AsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCys 1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 AACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGT 3420

Qy       1141 GluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeu 1160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 GAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTT 3480

Qy       1161 IleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIle 1180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3481 ATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATC 3540

Qy       1181 AsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSer 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3541 AATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCC 3600

Qy       1201 LysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGly 1220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3601 AAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGA 3660

Qy       1221 GluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsn 1240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3661 GAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAAT 3720

Qy       1241 GluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPhe 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 GAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTC 3780

Qy       1261 ThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuVal 1280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3781 ACCTGGCCCATCAAAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCTGGTA 3840

Qy       1281 GlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuLeuValAlaAlaValVal 1300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3841 CAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGTGGTT 3900

Qy       1301 TrpLysIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeuLeuArgGluMet 1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3901 TGGAAGATCAAACAAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGAGATG 3960

Qy       1321 GlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThrAspGluGlu 1340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3961 CAACAGATGGCCAGCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGAGGAG 4020

Qy       1341 ProProAspLeuIleGlyGlySerIleLysThrValProLysProIleAlaLeuGluPro 1360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4021 CCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGAGCCG 4080

Qy       1361 CysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGly 1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4081 TGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCTGGGT 4140

Qy       1381 GlyIleProProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSer 1400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4141 GGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACATTTCT 4200

Qy       1401 GlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAsnArgLysGlnGln 1420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4201 CAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCAGCAG 4260

Qy       1421 ProProAlaGlnProGlyThrCysIle 1429
             |||||||||||||||||||||||||||
Db       4261 CCCCCTGCACAGCCTGGGACCTGCATC 4287
```

RESULT 2
AAS72659
ID AAS72659 standard; cDNA; 8561 BP.
XX
AC AAS72659;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #8463.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG08472.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 8463; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this

```
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 8561 BP; 2246 A; 1983 C; 2065 G; 2267 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              0           Length:         8561
Score:                  1429.00     Matches:        1429
Percent Similarity:     100.00%     Conservative:   0
Best Local Similarity:  100.00%     Mismatches:     0
Query Match:            100.00%     Indels:         0
DB:                     5           Gaps:           0

US-09-787-097-12 (1-1429) x AAS72659 (1-8561)

Qy          1 MetValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 60

Qy         21 AlaLeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArg 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGGGACTGGGACGTGACCAGGGCTGGGAGG 120

Qy         41 ProGlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSerProProLeuArgProArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGGGGCTGGGGGCCGGGCTGCGCCTCCCGCGGCTGCTGTCTCCACCGCTGCGGCCACGG 180

Qy         61 LeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeuProCysGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGCTGCTGCTGCTGTTGTTGCTCCCGCCGCCGCTGTTGCTGCTGCTGCTGCCCTGTGAG 240

Qy         81 AlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCGAGGCCGCGGCGGCGGCGGCGGCGGTGTCGGGCTCAGCCGCAGCCGAGGCCAAGGAA 300

Qy        101 CysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCys 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGTGACCGGCCCTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGC 360

Qy        121 ProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGly 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGA 420

Qy        141 SerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrp 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGG 480

Qy        161 LeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAG 540

Qy        181 CysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAla 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCT 600

Qy        201 AlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGluValVal 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTT 660

Qy        221 AlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGly 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGA 720

Qy        241 PheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCys 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGT 780

Qy        261 LysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsnTrpLysGlyGlu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAGATCAGTAATAGCAGCGAAACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAA 840

Qy        281 AlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsn 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAAT 900

Qy        301 SerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerVal 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTT 960

Qy        321 ProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuPro 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCC 1020

Qy        341 ArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMet 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATG 1080

Qy        361 PheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeu 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTT 1140

Qy        381 ProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATAC 1200

Qy        401 LysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsnValThrAsnGlu 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AAGGATAAAATTTACATGTATGGAGGAAAAATTGATCCAACTGGGAATGTGACCAATGAG 1260

Qy        421 LeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGlu 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGGAG 1320

Qy        441 GlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValVal 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTC 1380

Qy        461 MetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyr 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATAT 1440

Qy        481 AspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGly 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGT 1500

Qy        501 TyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyr 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTAC 1560

Qy        521 LysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAsp 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGAT 1620

Qy        541 ThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaVal 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 ACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTG 1680

Qy        561 IleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSer 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 ATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGC 1740

Qy        581 HisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrp 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGG 1800

Qy        601 SerValLeuProArgProAspLeuHisHisAspValAsnArgPheGlyHisSerAlaVal 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 TCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTC 1860

Qy        621 LeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIle 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATC 1920

Qy        641 LeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAla 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      1921 CTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCA 1980
Qy       661 GlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeu 680
Db      1981 GGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTG 2040
Qy       681 AlaThrAspGluGlnGluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAsp 700
Db      2041 GCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGAC 2100
Qy       701 HisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCys 720
Db      2101 CATGACAGATGTGACCAGCACACAGATTGTTACAGCTGTACAGCCAACACCAATGACTGC 2160
Qy       721 HisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSer 740
Db      2161 CACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCC 2220
Qy       741 IlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThr 760
Db      2221 ATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCCATGTACTACTGTAACAAGAAGACC 2280
Qy       761 SerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCys 780
Db      2281 AGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGC 2340
Qy       781 IleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeu 800
Db      2341 ATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTG 2400
Qy       801 LysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsn 820
Db      2401 AAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAAT 2460
Qy       821 AlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArg 840
Db      2461 GCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGA 2520
Qy       841 IleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArg 860
Db      2521 ATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGG 2580
Qy       861 LysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeu 880
Db      2581 AAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTA 2640
Qy       881 GlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSer 900
Db      2641 CAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGT 2700
Qy       901 ThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArg 920
Db      2701 ACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGG 2760
Qy       921 ProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGly 940
Db      2761 CCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGA 2820
Qy       941 AspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAsp 960
Db      2821 GATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGAC 2880
Qy       961 SerAsnAlaTyrValAlaSerProProGlyGlnCysMetGluTrpTyrThrMetSer 980
Db      2881 TCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGC 2940
Qy       981 ThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPro 1000
Db      2941 ACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCA 3000
Qy      1001 GlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySer 1020
Db      3001 GGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCC 3060

Qy      1021 TyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnPro 1040
Db      3061 TATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCC 3120
Qy      1041 LeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCys 1060
Db      3121 CTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGT 3180
Qy      1061 ProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCys 1080
Db      3181 CCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGT 3240
Qy      1081 GluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspPro 1100
Db      3241 GAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCC 3300
Qy      1101 ThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThr 1120
Db      3301 ACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACC 3360
Qy      1121 AsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCys 1140
Db      3361 AACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGT 3420
Qy      1141 GluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeu 1160
Db      3421 GAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTT 3480
Qy      1161 IleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIle 1180
Db      3481 ATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATC 3540
Qy      1181 AsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSer 1200
Db      3541 AATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCC 3600
Qy      1201 LysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGly 1220
Db      3601 AAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGA 3660
Qy      1221 GluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsn 1240
Db      3661 GAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAAT 3720
Qy      1241 GluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPhe 1260
Db      3721 GAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTC 3780
Qy      1261 ThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuVal 1280
Db      3781 ACCTGGCCCATCAAAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCTGGTA 3840
Qy      1281 GlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuLeuValAlaAlaValVal 1300
Db      3841 CAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGTGGTT 3900
Qy      1301 TrpLysIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeuLeuArgGluMet 1320
Db      3901 TGGAAGATCAAACAAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGAGATG 3960
Qy      1321 GlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThrAspGluGlu 1340
Db      3961 CAACAGATGGCCAGCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGAGGAG 4020
Qy      1341 ProProAspLeuIleGlyGlySerIleLysThrValProLysProIleAlaLeuGluPro 1360
Db      4021 CCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGAGCCG 4080
Qy      1361 CysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGly 1380
Db      4081 TGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCTGGGT 4140
```

```
Qy       1381 GlyIleProProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSer 1400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4141 GGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACATTTCT 4200

Qy       1401 GlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAsnArgLysGlnGln 1420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4201 CAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCAGCAG 4260

Qy       1421 ProProAlaGlnProGlyThrCysIle 1429
             |||||||||||||||||||||||||||
Db       4261 CCCCCTGCACAGCCTGGGACCTGCATC 4287
```

RESULT 3
AAZ52277
ID AAZ52277 standard; cDNA; 3819 BP.
XX
AC AAZ52277;
XX
DT 18-JUL-2000 (first entry)
XX
DE Human soluble attractin-2 cDNA.
XX
KW Human; soluble attractin-2; immune response; macrophage; monocyte;
KW T cell; immunostimulant; immunosuppressed patient; cancer;
KW immunodeficiency syndrome; transplant; autoimmune disease; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1. .3819
FT /*tag= a
FT /product= "Soluble attractin-2"
XX
PN WO200015651-A1.
XX
PD 23-MAR-2000.
XX
PF 14-SEP-1999; 99WO-US020948.
XX
PR 14-SEP-1998; 98US-0100137P.
XX
PA (DAND) DANA FARBER CANCER INST INC.
XX
PI Duke-Cohan JS, Schlossman SF;
XX
DR WPI; 2000-271373/23.
DR P-PSDB; AAY70692.
XX
PT Isolated nucleic acids encoding human attractin polypeptides useful for
PT enhancing immune responses.
XX
PS Claim 3; Fig 12; 120pp; English.
XX
CC The patent discloses four forms of human attractin polypeptides which
CC enhance immune response by promoting macrophage and monocyte spreading in
CC the presence of T cells. These include soluble attractin-1 and -2 and
CC membrane attractin-1 and -2. These various forms of attractin are encoded
CC by alternatively spliced mRNA molecule transcribed from a single gene.
CC The present sequence is a cDNA encoding human soluble attractin-2.
CC Attractin-2 differs from attractin-1 in having a 74 amino acid insert in
CC the N-terminal. This sequence can be used to enhance immune response in
CC immunosuppressed patients such as those undergoing chemo- and radio-
CC therapy treatment for cancer or those suffering from common variable
CC immunodeficiency syndrome. The proteins may also be used to screen
CC modulators (agonists and antagonists) of immune response which may also
CC be used to regulate immune reactions. Attractin antibodies can be used to
CC inhibit immune response in transplant recipients or patients afflicted
CC with autoimmune disease
XX
SQ Sequence 3819 BP; 969 A; 899 C; 991 G; 960 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 3819

Score:	1267.00	Matches:	1267
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	88.66%	Indels:	0
DB:	3	Gaps:	0

US-09-787-097-12 (1-1429) x AAZ52277 (1-3819)

```
Qy          1 MetValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTGGCCGCAGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCA 60

Qy         21 AlaLeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArg 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCGCTCGCGGGCAGGAGCGGCGGGCCGCACTGGGACTGGGACGTGACCAGGGCTGGGAGG 120

Qy         41 ProGlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSerProProLeuArgProArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGGGGCTGGGGGCCGGGCTGCGCCTCCCGCGGCTGCTGTCTCCACCGCTGCGGCCACGG 180

Qy         61 LeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeuProCysGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGCTGCTGCTGCTGTTGTTGCTCCCGCCGCCGCTGTTGCTGCTGCTGCTGCCCTGTGAG 240

Qy         81 AlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCGAGGCCGCGGCGGCGGCGGCGGCGGCGGTGTCGGGCTCAGCCGCAGCCGAGGCCAAGGAA 300

Qy        101 CysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCys 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGTGACCGGCCCTGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGC 360

Qy        121 ProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGly 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCCGCCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGA 420

Qy        141 SerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrp 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGG 480

Qy        161 LeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAG 540

Qy        181 CysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAla 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCT 600

Qy        201 AlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGluValVal 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTT 660

Qy        221 AlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGly 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGA 720

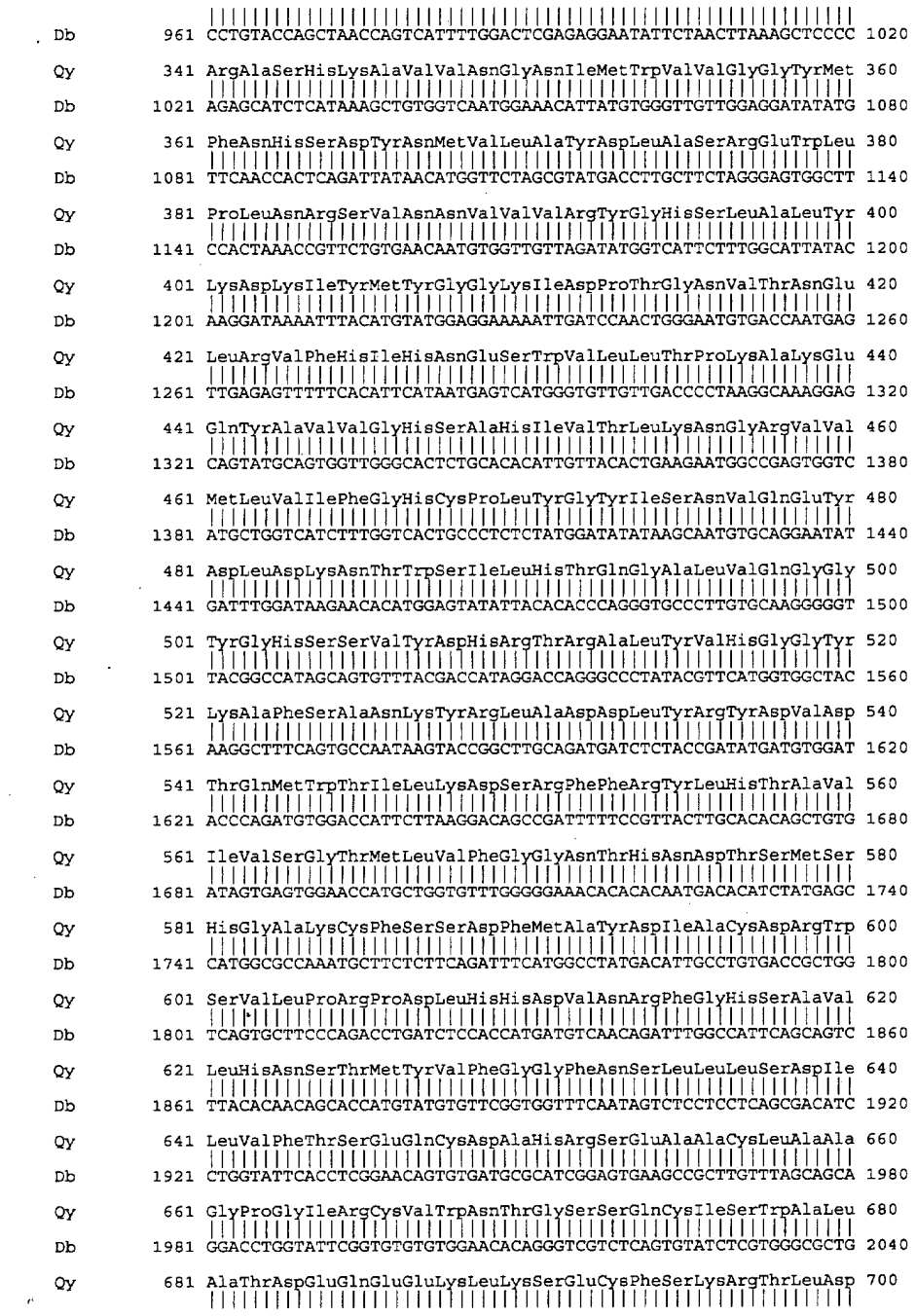
Qy        241 PheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCys 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGT 780

Qy        261 LysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsnTrpLysGlyGlu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAGATCAGTAATAGCAGCGAAACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAA 840

Qy        281 AlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsn 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAAT 900

Qy        301 SerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerVal 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTT 960

Qy        321 ProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuPro 340
```


Db 961 CCTGTACCAGCTAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCC 1020
Qy 341 ArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMet 360
Db 1021 AGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATG 1080
Qy 361 PheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeu 380
Db 1081 TTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTT 1140
Qy 381 ProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyr 400
Db 1141 CCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATAC 1200
Qy 401 LysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsnValThrAsnGlu 420
Db 1201 AAGGATAAAATTTACATGTATGGAGGAAAAATTGATCCAACTGGGAATGTGACCAATGAG 1260
Qy 421 LeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGlu 440
Db 1261 TTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGGAG 1320
Qy 441 GlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValVal 460
Db 1321 CAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTC 1380
Qy 461 MetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyr 480
Db 1381 ATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATAT 1440
Qy 481 AspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGly 500
Db 1441 GATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGT 1500
Qy 501 TyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyr 520
Db 1501 TACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTAC 1560
Qy 521 LysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAsp 540
Db 1561 AAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGAT 1620
Qy 541 ThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaVal 560
Db 1621 ACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTG 1680
Qy 561 IleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSer 580
Db 1681 ATAGTGAGTGGAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGC 1740
Qy 581 HisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrp 600
Db 1741 CATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGG 1800
Qy 601 SerValLeuProArgProAspLeuHisHisAspValAsnArgPheGlyHisSerAlaVal 620
Db 1801 TCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTC 1860
Qy 621 LeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIle 640
Db 1861 TTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATC 1920
Qy 641 LeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAla 660
Db 1921 CTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCA 1980
Qy 661 GlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeu 680
Db 1981 GGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTG 2040
Qy 681 AlaThrAspGluGlnGluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAsp 700
Db 2041 GCAACTGATGAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGAC 2100
Qy 701 HisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCys 720
Db 2101 CATGACAGATGTGACCAGCACACAGATTGTTACAGCTGTACAGCCAACACCAATGACTGC 2160
Qy 721 HisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSer 740
Db 2161 CACTGGTGCAATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCC 2220
Qy 741 IlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThr 760
Db 2221 ATTTTTAGGTATGAGAATTGCCCCAAGGATAACCCCATGTACTACTGTAACAAGAAGACC 2280
Qy 761 SerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCys 780
Db 2281 AGCTGCAGGAGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGC 2340
Qy 781 IleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeu 800
Db 2341 ATTGCCCTGCCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTG 2400
Qy 801 LysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsn 820
Db 2401 AAAATTACTACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAAT 2460
Qy 821 AlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArg 840
Db 2461 GCCCTTTTGGCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGA 2520
Qy 841 IleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArg 860
Db 2521 ATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGG 2580
Qy 861 LysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeu 880
Db 2581 AAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTA 2640
Qy 881 GlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSer 900
Db 2641 CAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGT 2700
Qy 901 ThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArg 920
Db 2701 ACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGG 2760
Qy 921 ProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGly 940
Db 2761 CCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGA 2820
Qy 941 AspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAsp 960
Db 2821 GATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGAC 2880
Qy 961 SerAsnAlaTyrValAlaSerProProPheGlyGlnCysMetGluTrpTyrThrMetSer 980
Db 2881 TCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGC 2940
Qy 981 ThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPro 1000
Db 2941 ACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCA 3000
Qy 1001 GlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySer 1020
Db 3001 GGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCC 3060
Qy 1021 TyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnPro 1040
Db 3061 TATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCC 3120
Qy 1041 LeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCys 1060
Db 3121 CTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGT 3180

```
Qy       1061 ProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCys 1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 CCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGT 3240

Qy       1081 GluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspPro 1100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 GAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCC 3300

Qy       1101 ThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThr 1120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 ACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACC 3360

Qy       1121 AsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCys 1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 AACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGT 3420

Qy       1141 GluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeu 1160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 GAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTT 3480

Qy       1161 IleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIle 1180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3481 ATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATC 3540

Qy       1181 AsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSer 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3541 AATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCC 3600

Qy       1201 LysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGly 1220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3601 AAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGA 3660

Qy       1221 GluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsn 1240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3661 GAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAAT 3720

Qy       1241 GluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPhe 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 GAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTC 3780

Qy       1261 ThrTrpProIleLysIleGln 1267
             |||||||||||||||||||||
Db       3781 ACCTGGCCCATCAAAATTCAG 3801
```

RESULT 4
AAZ91920
ID AAZ91920 standard; cDNA; 8590 BP.
XX
AC AAZ91920;
XX
DT 08-JUN-2000 (first entry)
XX
DE Human mahogany protein coding sequence #2.
XX
KW Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;
KW weight regulation; cell therapy; body weight disorder; cachexia;
KW anorexia; hyperpigmentation; increased metabolic rate disorder;
KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.
XX
OS Homo sapiens.
XX
PN WO200005373-A2.
XX
PD 03-FEB-2000.
XX
PF 21-JUL-1999; 99WO-US016484.
XX
PR 21-JUL-1998; 98US-0093630P.
PR 20-OCT-1998; 98US-0104978P.
PR 05-FEB-1999; 99US-00245041.
XX
PA (MILL-) MILLENIUM PHARM INC.
XX
PI Moore K, Nagle DL;
XX
DR WPI; 2000-195103/17.
DR P-PSDB; AAY81807.
XX
PT New human and murine mahogany genes, useful, e.g. for diagnosis and
PT treatment of body weight disorders.
XX
PS Claim 1; Fig 18a; 188pp; English.
XX
CC This sequence represents a human mahogany gene of the invention. The
CC mahogany genes are used: (i) to produce recombinant mahogany (mg)
CC proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming
CC therapeutics; (iii) as a source of diagnostic probes and primers for
CC detecting expression of mg genes or mutations, regulatory defects, in
CC this gene, or for isolation of related sequences; and (iv) in (cell-
CC based) gene therapy. (II) are used to raise specific antibodies (Ab); to
CC identify other (extra)cellular products involved in weight regulation,
CC and to screen for agents that disrupt interaction between (II) and other
CC macromolecules. The Ab are used to detect abnormal levels (or function)
CC of (II) (for diagnosis, prognosis or monitoring of treatment); to
CC evaluate (II)-expressing cells intended for cell therapy, and as
CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the
CC mg polypeptide) are used to identify agents (A) that modulate mg
CC activity. (A) are potentially useful for the treatment of body weight
CC disorders, particularly obesity, cachexia or anorexia, or other
CC conditions associated with the mg gene such as hyperpigmentation,
CC hyperphagia and disorders that result in increased metabolic rate
XX
SQ Sequence 8590 BP; 2275 A; 1990 C; 2061 G; 2264 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	8590
Score:	1141.00	Matches:	1341
Percent Similarity:	99.70%	Conservative:	0
Best Local Similarity:	99.70%	Mismatches:	2
Query Match:	79.85%	Indels:	4
DB:	3	Gaps:	0

US-09-787-097-12 (1-1429) x AAZ91920 (1-8590)

```
Qy         87 AlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGluCysAspArgProCysVal 106
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 GCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAAGGAATGTGACCGGCCGTGTGTC 315

Qy        107 AsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTrpValGly 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        316 AACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGC 375

Qy        127 GluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPheValThr 146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        376 GAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACA 435

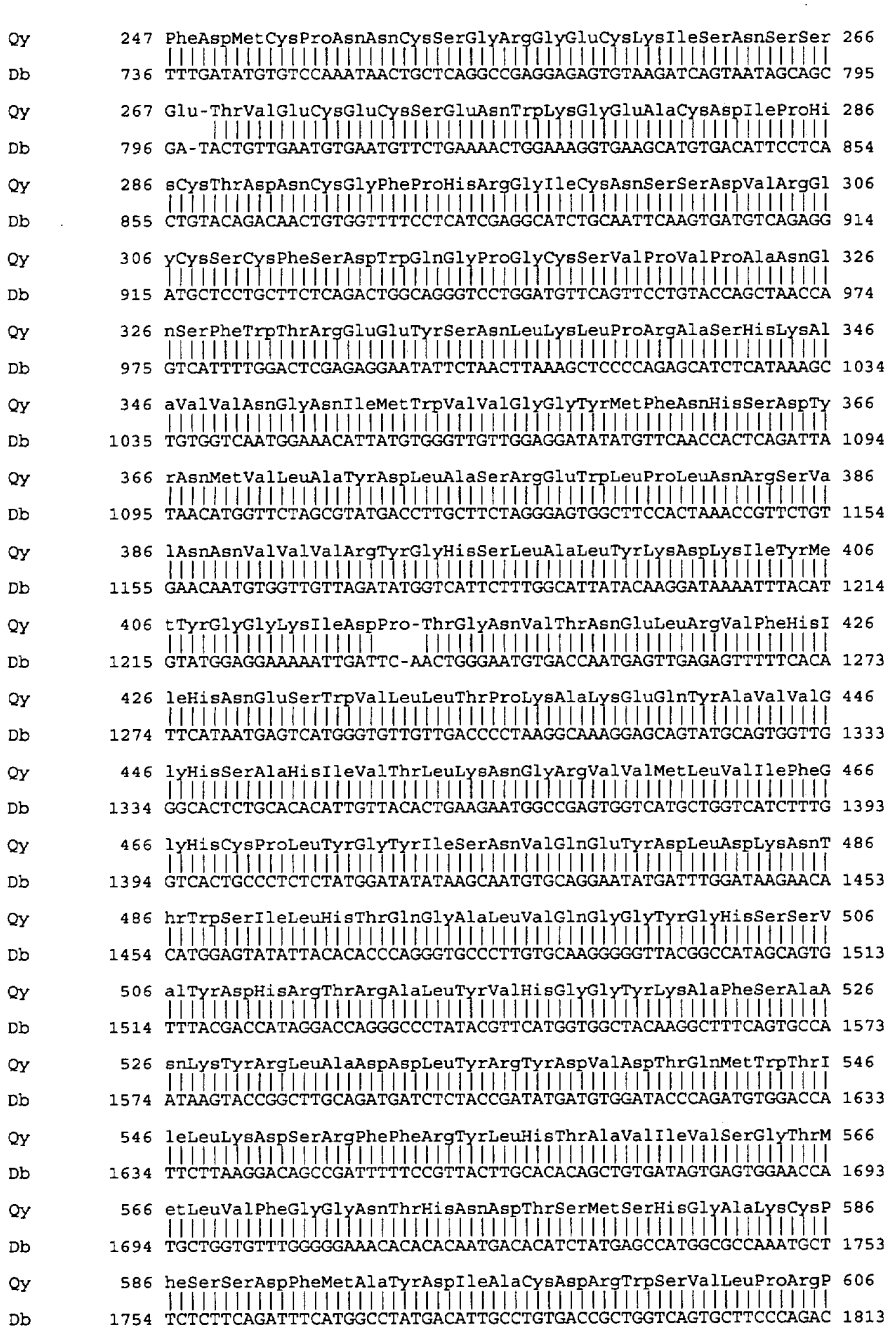
Qy        147 AspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeuIleGluGlyGlnPro 166
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        436 GATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCA 495

Qy        167 AsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLeu 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        496 AATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCATTTA 555

Qy        187 TyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaAlaPheSerGlyLeuIle 206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 TATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATT 615

Qy        207 ValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGlyTyrAla 226
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 GTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCC 675

Qy        227 LeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThrTyrSer 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 TTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTACAGT 735
```


Qy 247 PheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSerSer 266
Db 736 TTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGC 795
Qy 267 Glu-ThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIleProHi 286
Db 796 GA-TACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCA 854
Qy 286 sCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspValArgGl 306
Db 855 CTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGG 914
Qy 306 yCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAlaAsnGl 326
Db 915 ATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCA 974
Qy 326 nSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHisLysAl 346
Db 975 GTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGC 1034
Qy 346 aValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHisSerAspTy 366
Db 1035 TGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTA 1094
Qy 366 rAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSerVa 386
Db 1095 TAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGT 1154
Qy 386 lAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyrMe 406
Db 1155 GAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACAT 1214
Qy 406 tTyrGlyGlyLysIleAspPro-ThrGlyAsnValThrAsnGluLeuArgValPheHisI 426
Db 1215 GTATGGAGGAAAAATTGATTC-AACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACA 1273
Qy 426 leHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaValValG 446
Db 1274 TTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTG 1333
Qy 446 lyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValIlePheG 466
Db 1334 GGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTG 1393
Qy 466 lyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsnT 486
Db 1394 GTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACA 1453
Qy 486 hrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisSerServ 506
Db 1454 CATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTG 1513
Qy 506 alTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaA 526
Db 1514 TTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCCA 1573
Qy 526 snLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetTrpThrI 546
Db 1574 ATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCA 1633
Qy 546 leLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerGlyThrM 566
Db 1634 TTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACCA 1693
Qy 566 etLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysP 586
Db 1694 TGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATGCT 1753
Qy 586 heSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuProArgP 606
Db 1754 TCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGAC 1813
Qy 606 roAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrM 626
Db 1814 CTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACAACAGCACCA 1873
Qy 626 etTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSerG 646
Db 1874 TGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCGG 1933
Qy 646 luGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgC 666
Db 1934 AACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGT 1993
Qy 666 ysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnG 686
Db 1994 GTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAAG 2053
Qy 686 luGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArgCysAspG 706
Db 2054 AAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACC 2113
Qy 706 lnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAspH 726
Db 2114 AGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGCAATGACC 2173
Qy 726 isCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGluA 746
Db 2174 ATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGA 2233
Qy 746 snCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArgSerCysA 766
Db 2234 ATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTG 2293
Qy 766 laLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuProGluA 786
Db 2294 CCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAA 2353
Qy 786 snIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThrThrAlaL 806
Db 2354 ATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCA 2413
Qy 806 ysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSerL 826
Db 2414 AGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTC 2473
Qy 826 euThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGlnSerSerG 846
Db 2474 TTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTC 2533
Qy 846 lnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerT 866
Db 2534 AGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCT 2593
Qy 866 yrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerG 886
Db 2594 ACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTG 2653
Qy 886 luProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysA 906
Db 2654 AGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGG 2713
Qy 906 laAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSerA 926
Db 2714 CTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTG 2773
Qy 926 laLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGlyS 946
Db 2774 CTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGGCA 2833
Qy 946 erSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrValA 966
Db 2834 GCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGTGG 2893
Qy 966 laSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSerThrCysProProGluA 986

```
Db       2894 CCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGAAA 2953

Qy        986 snCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCysT 1006

Db       2954 ATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTA 3013

Qy       1006 hrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyProValL 1026

Db       3014 CTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGA 3073

Qy       1026 ysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerSerM 1046

Db       3074 AGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGCA 3133

Qy       1046 etCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnCysA 1066

Db       3134 TGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCA 3193

Qy       1066 snGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrThrG 1086

Db       3194 ACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCACAG 3253

Qy       1086 lyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyGlyLysC 1106

Db       3254 GCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAAT 3313

Qy       1106 ysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysP 1126

Db       3314 GTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGCT 3373

Qy       1126 heCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGluAsnArgT 1146

Db       3374 TCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGAT 3433

Qy       1146 yrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrGlnPheT 1166

Db       3434 ACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTTCA 3493

Qy       1166 hrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIleAsnPheValAlaThrP 1186

Db       3494 CCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTC 3553

Qy       1186 roAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeuA 1206

Db       3554 CTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCTCA 3613

Qy       1206 snIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProValV 1226

Db       3614 ACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTG 3673

Qy       1226 alSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPheAspPheA 1246

Db       3674 TTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTC 3733

Qy       1246 rgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpProIleLysI 1266

Db       3734 GCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAAAA 3793

Qy       1266 leGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePheValThrP 1286

Db       3794 TTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCTGGTACAGTTCTTCGTGACTT 3853

Qy       1286 hePheSerCysPheLeuSerLeuLeuLeuValAlaAlaValValTrpLysIleLysGlnS 1306

Db       3854 TCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGTGGTTTGGAAGATCAAACAAA 3913

Qy       1306 erCysTrpAlaSerArgArgArgGluGlnLeuLeuArgGluMetGlnGlnMetAlaSerA 1326

Db       3914 GTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGAGATGCAACAGATGGCCAGCC 3973

Qy       1326 rgProPheAlaSerValAsnValAlaLeuGluThrAspGluGluProProAspLeuIleG 1346

Db       3974 GTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTATTG 4033

Qy       1346 lyGlySerIleLysThrValProLysProIleAlaLeuGluProCysPheGlyAsnLysA 1366

Db       4034 GGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGAGCCGTGTTTTGGCAACAAAG 4093

Qy       1366 laAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyGlyIleProProProG 1386

Db       4094 CCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCTGGGTGGCATCCCTCCTCCTG 4153

Qy       1386 lyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetProIleV 1406

Db       4154 GGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGATAG 4213

Qy       1406 alTyrLysGluLysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnProG 1426

Db       4214 TGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGCCTG 4273

Qy       1426 lyThrCysIle 1429

Db       4274 GGACCTGCATC 4284
```

RESULT 5
AAZ52275
ID AAZ52275 standard; cDNA; 4068 BP.
XX
AC AAZ52275;
XX
DT 18-JUL-2000 (first entry)
XX
DE Human membrane attractin-1 cDNA.
XX
KW Human; membrane attractin-1; immune response; macrophage; monocyte;
KW T cell; immunostimulant; immunosuppressed patient; cancer;
KW immunodeficiency syndrome; transplant; autoimmune disease; ss.
XX
OS Homo sapiens.
XX
FH Key   Location/Qualifiers
FT CDS   1. .4068
FT   /*tag= a
FT   /product= "Membrane attractin-1"
XX
PN WO200015651-A1.
XX
PD 23-MAR-2000.
XX
PF 14-SEP-1999; 99WO-US020948.
XX
PR 14-SEP-1998; 98US-0100137P.
XX
PA (DAND) DANA FARBER CANCER INST INC.
XX
PI Duke-Cohan JS, Schlossman SF;
XX
DR WPI; 2000-271373/23.
DR P-PSDB; AAY70690.
XX
PT Isolated nucleic acids encoding human attractin polypeptides useful for
PT enhancing immune responses.
XX
PS Claim 3; Fig 10; 120pp; English.
XX
CC The patent discloses four forms of human attractin polypeptides which
CC enhance immune response by promoting macrophage and monocyte spreading in
CC the presence of T cells. These include soluble attractin-1 and -2 and
CC membrane attractin-1 and -2. These various forms of attractin are encoded
CC by alternatively spliced mRNA molecule transcribed from a single gene.
CC The present sequence is a cDNA encoding human membrane attractin-1.
CC Membrane attractin differs from soluble attractin in having a
CC transmembrane domain and a cytoplasmic domain at the C-terminal. This
CC sequence can be used to enhance immune response in immunosuppressed
CC patients such as those undergoing chemo- and radio-therapy treatment for

CC cancer or those suffering from common variable immunodeficiency syndrome.
CC The proteins may also be used to screen modulators (agonists and
CC antagonists) of immune response which may also be used to regulate immune
CC reactions. Attractin antibodies can be used to inhibit immune response in
CC transplant recipients or patients afflicted with autoimmune disease
XX
SQ Sequence 4068 BP; 1042 A; 948 C; 1027 G; 1051 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	4068
Score:	1123.00	Matches:	1323
Percent Similarity:	99.70%	Conservative:	0
Best Local Similarity:	99.70%	Mismatches:	2
Query Match:	78.59%	Indels:	4
DB:	3	Gaps:	0

US-09-787-097-12 (1-1429) x AAZ52275 (1-4068)

```
Qy        105 CysValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTrp 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 TGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGG 150

Qy        125 ValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPhe 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 GTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTT 210

Qy        145 ValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeuIleGluGly 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 GTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGA 270

Qy        165 GlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAsp 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 CAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGAC 330

Qy        185 HisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaAlaPheSerGly 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 CATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGC 390

Qy        205 LeuIleValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGly 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 CTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGT 450

Qy        225 TyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 TATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACT 510

Qy        245 TyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsn 264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 TACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAAT 570

Qy        265 SerSerGlu-ThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIl 284
              ||||||    ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        571 AGCAGCGA-TACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACAT 629

Qy        284 eProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspVa 304
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        630 TCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGT 689

Qy        304 lArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAl 324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        690 CAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGC 749

Qy        324 aAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHi 344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        750 TAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCA 809

Qy        344 sLysAlaValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHisSe 364
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        810 TAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTC 869

Qy        364 rAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnAr 384
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        870 AGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCG 929

Qy        384 gSerValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIl 404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        930 TTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAAT 989

Qy        404 eTyrMetTyrGlyGlyLysIleAspPro-ThrGlyAsnValThrAsnGluLeuArgValP 424
              |||||||||||||||||||||||||    ||||||||||||||||||||||||||||||
Db        990 TTACATGTATGGAGGAAAAATTGATTC-AACTGGGAATGTGACCAATGAGTTGAGAGTTT 1048

Qy        424 heHisIleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaV 444
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1049 TTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAG 1108

Qy        444 alValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValI 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1109 TGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCA 1168

Qy        464 lePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspL 484
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1169 TCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATA 1228

Qy        484 ysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisS 504
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1229 AGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATA 1288

Qy        504 erSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheS 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1289 GCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCA 1348

Qy        524 erAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetT 544
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1349 GTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGT 1408

Qy        544 rpThrIleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerG 564
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1409 GGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTG 1468

Qy        564 lyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaL 584
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1469 GAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGCCATGGCGCCA 1528

Qy        584 ysCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuP 604
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1529 AATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTC 1588

Qy        604 roArgProAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnS 624
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1589 CCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACAACA 1648

Qy        624 erThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPheT 644
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1649 GCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCA 1708

Qy        644 hrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyI 664
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1709 CCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTA 1768

Qy        664 leArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspG 684
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1769 TTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATG 1828

Qy        684 luGlnGluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArgC 704
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1829 AACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGAT 1888

Qy        704 ysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCysA 724
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1889 GTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGCA 1948

Qy        724 snAspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgT 744
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1949 ATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGT 2008
```

```
Qy       744 yrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArgS 764
Db      2009 ATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGGA 2068

Qy       764 erCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuP 784
Db      2069 GCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGC 2128

Qy       784 roGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThrT 804
Db      2129 CCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTA 2188

Qy       804 hrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuA 824
Db      2189 CTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGG 2248

Qy       824 laSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGlnS 844
Db      2249 CTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGT 2308

Qy       844 erSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnV 864
Db      2309 CATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATG 2368

Qy       864 alSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetP 884
Db      2369 TGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGC 2428

Qy       884 roSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyL 904
Db      2429 CGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGAC 2488

Qy       904 euLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnH 924
Db      2489 TGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACC 2548

Qy       924 isSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThrS 944
Db      2549 ACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCA 2608

Qy       944 erGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaT 964
Db      2609 GCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCT 2668

Qy       964 yrValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSerThrCysProP 984
Db      2669 ATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCC 2728

Qy       984 roGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyT 1004
Db      2729 CTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCT 2788

Qy      1004 rpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyP 1024
Db      2789 GGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGAC 2848

Qy      1024 roValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnS 1044
Db      2849 CAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATT 2908

Qy      1044 erSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysG 1064
Db      2909 CCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCC 2968

Qy      1064 lnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuT 1084
Db      2969 AATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGA 3028

Qy      1084 hrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyG 1104
Db      3029 CCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAG 3088

Qy      1104 lyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyL 1124
Db      3089 GGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCA 3148

Qy      1124 ysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGluA 1144
Db      3149 AGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAA 3208

Qy      1144 snArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrG 1164
Db      3209 ATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATC 3268

Qy      1164 lnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIleAsnPheValA 1184
Db      3269 AGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGG 3328

Qy      1184 laThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheA 1204
Db      3329 CTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCA 3388

Qy      1204 snLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMetP 1224
Db      3389 ACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGC 3448

Qy      1224 roValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPheA 1244
Db      3449 CTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTG 3508

Qy      1244 spPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpProI 1264
Db      3509 ATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCA 3568

Qy      1264 leLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePheV 1284
Db      3569 TCAAAATTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCTGGTACAGTTCTTCG 3628

Qy      1284 alThrPhePheSerCysPheLeuSerLeuLeuLeuValAlaAlaValValTrpLysIleL 1304
Db      3629 TGACTTTCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGTGGTTTGGAAGATCA 3688

Qy      1304 ysGlnSerCysTrpAlaSerArgArgArgGluGlnLeuLeuArgGluMetGlnGlnMetA 1324
Db      3689 AACAAAGTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGAGATGCAACAGATGG 3748

Qy      1324 laSerArgProPheAlaSerValAsnValAlaLeuGluThrAspGluGluProProAspL 1344
Db      3749 CCAGCCGTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGAGGAGCCTCCTGATC 3808

Qy      1344 euIleGlyGlySerIleLysThrValProLysProIleAlaLeuGluProCysPheGlyA 1364
Db      3809 TTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGAGCCGTGTTTTGGCA 3868

Qy      1364 snLysAlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyGlyIleProP 1384
Db      3869 ACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCTGGGTGGCATCCCTC 3928

Qy      1384 roProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetP 1404
Db      3929 CTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGC 3988

Qy      1404 roIleValTyrLysGluLysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaG 1424
Db      3989 CGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCAGCAGCCCCCTGCAC 4048

Qy      1424 lnProGlyThrCysIle 1429
Db      4049 AGCCTGGGACCTGCATC 4065
```

RESULT 6
AAZ91921
ID AAZ91921 standard; cDNA; 4072 BP.
XX
AC AAZ91921;
XX

DT 08-JUN-2000 (first entry)
XX
DE Human mahogany protein coding sequence #3.
XX
KW Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;
KW weight regulation; cell therapy; body weight disorder; cachexia;
KW anorexia; hyperpigmentation; increased metabolic rate disorder;
KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.
XX
OS Homo sapiens.
XX
PN WO200005373-A2.
XX
PD 03-FEB-2000.
XX
PF 21-JUL-1999; 99WO-US016484.
XX
PR 21-JUL-1998; 98US-0093630P.
PR 20-OCT-1998; 98US-0104978P.
PR 05-FEB-1999; 99US-00245041.
XX
PA (MILL-) MILLENIUM PHARM INC.
XX
PI Moore K, Nagle DL;
XX
DR WPI; 2000-195103/17.
DR P-PSDB; AAY81808.
XX
PT New human and murine mahogany genes, useful, e.g. for diagnosis and
PT treatment of body weight disorders.
XX
PS Claim 1; Fig 19a; 188pp; English.
XX
CC This sequence represents a human mahogany gene of the invention. The
CC mahogany genes are used: (i) to produce recombinant mahogany (mg)
CC proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming
CC therapeutics; (iii) as a source of diagnostic probes and primers for
CC detecting expression of mg genes or mutations, regulatory defects, in
CC this gene, or for isolation of related sequences; and (iv) in (cell-
CC based) gene therapy. (II) are used to raise specific antibodies (Ab); to
CC identify other (extra)cellular products involved in weight regulation,
CC and to screen for agents that disrupt interaction between (II) and other
CC macromolecules. The Ab are used to detect abnormal levels (or function)
CC of (II) (for diagnosis, prognosis or monitoring of treatment); to
CC evaluate (II)-expressing cells intended for cell therapy, and as
CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the
CC mg polypeptide) are used to identify agents (A) that modulate mg
CC activity. (A) are potentially useful for the treatment of body weight
CC disorders, particularly obesity, cachexia or anorexia, or other
CC conditions associated with the mg gene such as hyperpigmentation,
CC hyperphagia and disorders that result in increased metabolic rate
XX
SQ Sequence 4072 BP; 1087 A; 940 C; 1045 G; 1000 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	4072
Score:	979.00	Matches:	1179
Percent Similarity:	99.66%	Conservative:	0
Best Local Similarity:	99.66%	Mismatches:	2
Query Match:	68.51%	Indels:	4
DB:	3	Gaps:	0

US-09-787-097-12 (1-1429) x AAZ91921 (1-4072)

```
Qy        87 AlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGluCysAspArgProCysVal 106
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       256 GCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAAGGAATGTGACCGGCCGTGTGTC 315

Qy       107 AsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTrpValGly 126
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       316 AACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGC 375

Qy       127 GluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPheValThr 146
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       376 GAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACA 435

Qy       147 AspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeuIleGluGlyGlnPro 166
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       436 GATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCA 495

Qy       167 AsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLeu 186
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       496 AATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCATTTA 555

Qy       187 TyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaAlaPheSerGlyLeuIle 206
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       556 TATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATT 615

Qy       207 ValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGlyTyrAla 226
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       616 GTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCC 675

Qy       227 LeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThrTyrSer 246
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       676 TTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTACAGT 735

Qy       247 PheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSerSer 266
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       736 TTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGC 795

Qy       267 Glu-ThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIleProHi 286
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       796 GA-TACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCA 854

Qy       286 sCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspValArgGl 306
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       855 CTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGG 914

Qy       306 yCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAlaAsnGl 326
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       915 ATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCA 974

Qy       326 nSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHisLysAl 346
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       975 GTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGC 1034

Qy       346 aValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHisSerAspTy 366
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1035 TGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTA 1094

Qy       366 rAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSerVa 386
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1095 TAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGT 1154

Qy       386 lAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyrMe 406
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1155 GAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACAT 1214

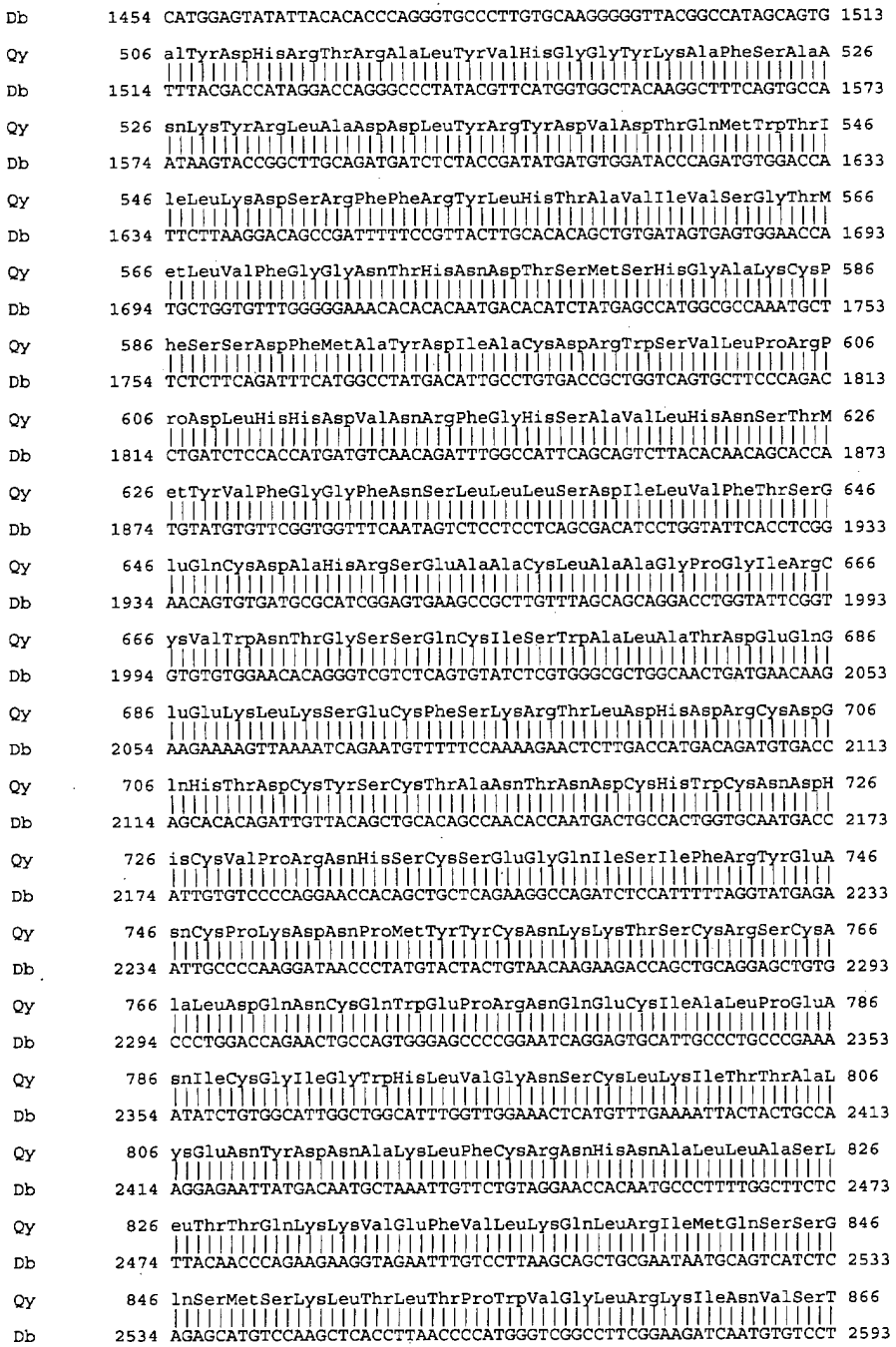
Qy       406 tTyrGlyGlyLysIleAspPro-ThrGlyAsnValThrAsnGluLeuArgValPheHisI 426
             ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db      1215 GTATGGAGGAAAAATTGATTC-AACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACA 1273

Qy       426 leHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaValValG 446
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1274 TTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTG 1333

Qy       446 lyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValIlePheG 466
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1334 GGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTG 1393

Qy       466 lyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsnT 486
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1394 GTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACA 1453

Qy       486 hrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisSerServ 506
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```


Db 1454 CATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTG 1513
Qy 506 alTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaA 526
Db 1514 TTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCCA 1573
Qy 526 snLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetTrpThrI 546
Db 1574 ATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCA 1633
Qy 546 leLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerGlyThrM 566
Db 1634 TTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACCA 1693
Qy 566 etLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysP 586
Db 1694 TGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATGCT 1753
Qy 586 heSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuProArgP 606
Db 1754 TCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGAC 1813
Qy 606 roAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrM 626
Db 1814 CTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACAACAGCACCA 1873
Qy 626 etTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSerG 646
Db 1874 TGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCGG 1933
Qy 646 luGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgC 666
Db 1934 AACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGT 1993
Qy 666 ysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnG 686
Db 1994 GTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAAG 2053
Qy 686 luGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArgCysAspG 706
Db 2054 AAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACC 2113
Qy 706 lnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAspH 726
Db 2114 AGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGCAATGACC 2173
Qy 726 isCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGluA 746
Db 2174 ATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGA 2233
Qy 746 snCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArgSerCysA 766
Db 2234 ATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTG 2293
Qy 766 laLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuProGluA 786
Db 2294 CCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAA 2353
Qy 786 snIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThrThrAlaL 806
Db 2354 ATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCA 2413
Qy 806 ysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSerL 826
Db 2414 AGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTC 2473
Qy 826 euThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGlnSerSerG 846
Db 2474 TTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTC 2533
Qy 846 lnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerT 866
Db 2534 AGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCT 2593
Qy 866 yrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerG 886
Db 2594 ACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTG 2653
Qy 886 luProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysA 906
Db 2654 AGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGG 2713
Qy 906 laAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSerA 926
Db 2714 CTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTG 2773
Qy 926 laLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGlyS 946
Db 2774 CTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGGCA 2833
Qy 946 erSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrValA 966
Db 2834 GCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGTGG 2893
Qy 966 laSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSerThrCysProProGluA 986
Db 2894 CCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGAAA 2953
Qy 986 snCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCysT 1006
Db 2954 ATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTA 3013
Qy 1006 hrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyProValL 1026
Db 3014 CTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGA 3073
Qy 1026 ysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerSerM 1046
Db 3074 AGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGCA 3133
Qy 1046 etCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnCysA 1066
Db 3134 TGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCA 3193
Qy 1066 snGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrThrG 1086
Db 3194 ACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCACAG 3253
Qy 1086 lyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyGlyLysC 1106
Db 3254 GCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAAT 3313
Qy 1106 ysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysP 1126
Db 3314 GTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGCT 3373
Qy 1126 heCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGluAsnArgT 1146
Db 3374 TCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGAT 3433
Qy 1146 yrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrGlnPheT 1166
Db 3434 ACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTTCA 3493
Qy 1166 hrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIleAsnPheValAlaThrP 1186
Db 3494 CCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTC 3553
Qy 1186 roAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeuA 1206
Db 3554 CTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCTCA 3613
Qy 1206 snIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProValV 1226
Db 3614 ACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTG 3673

```
Qy       1226 alSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPheAspPheA 1246
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3674 TTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTC 3733

Qy       1246 rgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpProIleLysI 1266
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3734 GCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAAAA 3793

Qy       1266 leGln 1267
             |||||
Db       3794 TTCAG 3798

RESULT 7
AAZ52274
ID   AAZ52274 standard; cDNA; 3597 BP.
XX
AC   AAZ52274;
XX
DT   18-JUL-2000  (first entry)
XX
DE   Human soluble attractin-1 cDNA.
XX
KW   Human; soluble attractin-1; immune response; macrophage; monocyte;
KW   T cell; immunostimulant; immunosuppressed patient; cancer;
KW   immunodeficiency syndrome; transplant; autoimmune disease; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..  .3597
FT                   /*tag=  a
FT                   /product= "Soluble attractin-1"
XX
PN   WO200015651-A1.
XX
PD   23-MAR-2000.
XX
PF   14-SEP-1999;   99WO-US020948.
XX
PR   14-SEP-1998;   98US-0100137P.
XX
PA   (DAND ) DANA FARBER CANCER INST INC.
XX
PI   Duke-Cohan JS,  Schlossman SF;
XX
DR   WPI; 2000-271373/23.
DR   P-PSDB; AAY70689.
XX
PT   Isolated nucleic acids encoding human attractin polypeptides useful for
PT   enhancing immune responses.
XX
PS   Claim 3; Fig 8; 120pp; English.
XX
CC   The patent discloses four forms of human attractin polypeptides which
CC   enhance immune response by promoting macrophage and monocyte spreading in
CC   the presence of T cells. These include soluble attractin-1 and -2 and
CC   membrane attractin-1 and -2. These various forms of attractin are encoded
CC   by alternatively spliced mRNA molecule transcribed from a single gene.
CC   The present sequence is a cDNA encoding soluble attractin-1 obtained from
CC   activated T cell and human foetal liver libraries. This sequence can be
CC   used to enhance immune response in immunosuppressed patients such as
CC   those undergoing chemo- and radio-therapy treatment for cancer or those
CC   suffering from common variable immunodeficiency syndrome. The proteins
CC   may also be used to screen modulators (agonists and antagonists) of
CC   immune responses which may also be used to regulate immune reactions.
CC   Attractin antibodies can be used to inhibit immune response in transplant
CC   recipients or patients afflicted with autoimmune disease
XX
SQ   Sequence 3597 BP; 951 A; 822 C; 898 G; 926 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              0           Length:        3597
Score:                  874.00      Matches:       1160
Percent Similarity:     99.49%      Conservative:  0
Best Local Similarity:  99.49%      Mismatches:    3
Query Match:            61.16%      Indels:        6
DB:                     3           Gaps:          0

US-09-787-097-12 (1-1429) x AAZ52274 (1-3597)

Qy        105 CysValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTrp 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 TGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGG 150

Qy        125 ValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPhe 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 GTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTT 210

Qy        145 ValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeuIleGluGly 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 GTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGA 270

Qy        165 GlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAsp 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 CAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGAC 330

Qy        185 HisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaAlaPheSerGly 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 CATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGC 390

Qy        205 LeuIleValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGly 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 CTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGT 450

Qy        225 TyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 TATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACT 510

Qy        245 TyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsn 264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 TACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAAT 570

Qy        265 SerSerGlu-ThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIl 284
              ||||||    ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        571 AGCAGCGA-TACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACAT 629

Qy        284 eProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspVa 304
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        630 TCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGT 689

Qy        304 lArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAl 324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        690 CAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGC 749

Qy        324 aAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHi 344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        750 TAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCA 809

Qy        344 sLysAlaValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHisSe 364
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        810 TAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTC 869

Qy        364 rAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnAr 384
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        870 AGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCG 929

Qy        384 gSerValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIl 404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        930 TTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAAT 989

Qy        404 eTyrMetTyrGlyGlyLysIleAspPro-ThrGlyAsnValThrAsnGluLeuArgValP 424
              ||||||||||||||||||||||||    ||||||||||||||||||||||||||||||||
Db        990 TTACATGTATGGAGGAAAAATTGATTC-AACTGGGAATGTGACCAATGAGTTGAGAGTTT 1048

Qy        424 heHisIleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaV 444
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1049 TTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAG 1108

Qy        444 alValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValI 464

Db       1109 TGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCA 1168

Qy        464 lePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspL 484

Db       1169 TCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATA 1228

Qy        484 ysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisS 504

Db       1229 AGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATA 1288

Qy        504 erSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheS 524

Db       1289 GCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCA 1348

Qy        524 erAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetT 544

Db       1349 GTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGT 1408

Qy        544 rpThrIleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerG 564

Db       1409 GGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTG 1468

Qy        564 lyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaL 584

Db       1469 GAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGCCATGGCGCCA 1528

Qy        584 ysCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuP 604

Db       1529 AATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTC 1588

Qy        604 roArgProAspLeuHisHisAspValAsnArgPheGlyHisSer-AlaValLeuHisAsn 623

Db       1589 CCAGACCTGATTC-CACCATGATGTCAACAGATTTGGCCATTCCAGCAGTCTTACACAAC 1647

Qy        624 SerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPhe 643

Db       1648 AGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTC 1707

Qy        644 ThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGly 663

Db       1708 ACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGT 1767

Qy        664 IleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAsp 683

Db       1768 ATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGAT 1827

Qy        684 GluGlnGluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArg 703

Db       1828 GAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGA 1887

Qy        704 CysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCys 723

Db       1888 TGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGC 1947

Qy        724 AsnAspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArg 743

Db       1948 AATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGG 2007

Qy        744 TyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArg 763

Db       2008 TATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGG 2067

Qy        764 SerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeu 783

Db       2068 AGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTG 2127

Qy        784 ProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThr 803

Db       2128 CCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACT 2187

Qy        804 ThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeu 823

Db       2188 ACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTG 2247

Qy        824 AlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGln 843

Db       2248 GCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAG 2307

Qy        844 SerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsn 863

Db       2308 TCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAAT 2367

Qy        864 ValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMet 883

Db       2368 GTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATG 2427

Qy        884 ProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGly 903

Db       2428 CCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGA 2487

Qy        904 LeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsn 923

Db       2488 CTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAAC 2547

Qy        924 HisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThr 943

Db       2548 CACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACC 2607

Qy        944 SerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAla 963

Db       2608 AGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCC 2667

Qy        964 TyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSerThrCysPro 983

Db       2668 TATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCC 2727

Qy        984 ProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGly 1003

Db       2728 CCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGC 2787

Qy       1004 TrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGly 1023

Db       2788 TGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGA 2847

Qy       1024 ProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsn 1043

Db       2848 CCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAAT 2907

Qy       1044 SerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCys 1063

Db       2908 TCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGC 2967

Qy       1064 GlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeu 1083

Db       2968 CAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTG 3027

Qy       1084 ThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGly 1103

Db       3028 ACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGA 3087

Qy       1104 GlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGly 1123

Db       3088 GGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGC 3147

Qy       1124 LysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGlu 1143

Db       3148 AAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAA 3207

Qy       1144 AsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyr 1163

Db       3208 AATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTAT 3267
```

```
Qy       1164 GlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIleAsnPheVal 1183
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3268 CAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTG 3327

Qy       1184 AlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPhe 1203
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3328 GCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTC 3387

Qy       1204 AsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMet 1223
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3388 AACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATG 3447

Qy       1224 ProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPhe 1243
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3448 CCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTT 3507

Qy       1244 AspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpPro 1263
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3508 GATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCC 3567

Qy       1264 IleLysIleGln 1267
             ||||||||||||
Db       3568 ATCAAAATTCAG 3579
```

RESULT 8
AAS72657
ID AAS72657 standard; cDNA; 3597 BP.
XX
AC AAS72657;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #8461.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG08470.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 8461; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 3597 BP; 951 A; 822 C; 898 G; 926 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	3597
Score:	874.00	Matches:	1160
Percent Similarity:	99.49%	Conservative:	0
Best Local Similarity:	99.49%	Mismatches:	3
Query Match:	61.16%	Indels:	6
DB:	5	Gaps:	0

US-09-787-097-12 (1-1429) x AAS72657 (1-3597)

```
Qy        105 CysValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTrp 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 TGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGG 150

Qy        125 ValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPhe 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 GTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTT 210

Qy        145 ValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeuIleGluGly 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 GTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGA 270

Qy        165 GlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAsp 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 CAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGAC 330

Qy        185 HisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaAlaPheSerGly 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 CATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGC 390

Qy        205 LeuIleValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGly 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 CTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGT 450

Qy        225 TyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 TATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACT 510

Qy        245 TyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsn 264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 TACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAAT 570

Qy        265 SerSerGlu-ThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIl 284
              |||||||    ||||||||||||||||||||||||||||||||||||||||||||||||
Db        571 AGCAGCGA-TACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACAT 629

Qy        284 eProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspVa 304
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        630 TCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGT 689

Qy        304 lArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAl 324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        690 CAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGC 749

Qy        324 aAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHi 344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        750 TAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCA 809

Qy        344 sLysAlaValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHisSe 364
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        810 TAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTC 869
Qy        364 rAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnAr 384
Db        870 AGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCG 929
Qy        384 gSerValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIl 404
Db        930 TTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAAT 989
Qy        404 eTyrMetTyrGlyGlyLysIleAspPro-ThrGlyAsnValThrAsnGluLeuArgValP 424
Db        990 TTACATGTATGGAGGAAAAATTGATTC-AACTGGGAATGTGACCAATGAGTTGAGAGTTT 1048
Qy        424 heHisIleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaV 444
Db       1049 TTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAG 1108
Qy        444 alValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValI 464
Db       1109 TGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCA 1168
Qy        464 lePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspL 484
Db       1169 TCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATA 1228
Qy        484 ysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisS 504
Db       1229 AGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATA 1288
Qy        504 erSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheS 524
Db       1289 GCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCA 1348
Qy        524 erAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetT 544
Db       1349 GTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGT 1408
Qy        544 rpThrIleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerG 564
Db       1409 GGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTG 1468
Qy        564 lyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaL 584
Db       1469 GAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGCCATGGCGCCA 1528
Qy        584 ysCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuP 604
Db       1529 AATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTC 1588
Qy        604 roArgProAspLeuHisHisAspValAsnArgPheGlyHisSer-AlaValLeuHisAsn 623
Db       1589 CCAGACCTGATTC-CACCATGATGTCAACAGATTTGGCCATTCCAGCAGTCTTACACAAC 1647
Qy        624 SerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPhe 643
Db       1648 AGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTC 1707
Qy        644 ThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGly 663
Db       1708 ACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGT 1767
Qy        664 IleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAsp 683
Db       1768 ATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGAT 1827
Qy        684 GluGlnGluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArg 703
Db       1828 GAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGA 1887
Qy        704 CysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCys 723
Db       1888 TGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGC 1947

Qy        724 AsnAspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArg 743
Db       1948 AATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGG 2007
Qy        744 TyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArg 763
Db       2008 TATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGG 2067
Qy        764 SerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeu 783
Db       2068 AGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTG 2127
Qy        784 ProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThr 803
Db       2128 CCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACT 2187
Qy        804 ThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeu 823
Db       2188 ACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTG 2247
Qy        824 AlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGln 843
Db       2248 GCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAG 2307
Qy        844 SerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsn 863
Db       2308 TCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAAT 2367
Qy        864 ValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMet 883
Db       2368 GTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATG 2427
Qy        884 ProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGly 903
Db       2428 CCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGA 2487
Qy        904 LeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsn 923
Db       2488 CTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAAC 2547
Qy        924 HisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThr 943
Db       2548 CACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACC 2607
Qy        944 SerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAla 963
Db       2608 AGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCC 2667
Qy        964 TyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSerThrCysPro 983
Db       2668 TATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCC 2727
Qy        984 ProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGly 1003
Db       2728 CCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGC 2787
Qy       1004 TrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGly 1023
Db       2788 TGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGA 2847
Qy       1024 ProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsn 1043
Db       2848 CCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAAT 2907
Qy       1044 SerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCys 1063
Db       2908 TCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGC 2967
Qy       1064 GlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeu 1083
Db       2968 CAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTG 3027
```

```
Qy       1084 ThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGly 1103
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3028 ACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGA 3087

Qy       1104 GlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGly 1123
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3088 GGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGC 3147

Qy       1124 LysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGlu 1143
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3148 AAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAA 3207

Qy       1144 AsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyr 1163
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3208 AATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTAT 3267

Qy       1164 GlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIleAsnPheVal 1183
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3268 CAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTG 3327

Qy       1184 AlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPhe 1203
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3328 GCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTC 3387

Qy       1204 AsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMet 1223
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3388 AACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATG 3447

Qy       1224 ProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPhe 1243
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3448 CCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTT 3507

Qy       1244 AspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpPro 1263
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3508 GATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCC 3567

Qy       1264 IleLysIleGln 1267
             ||||||||||||
Db       3568 ATCAAAATTCAG 3579
```

RESULT 9
ADD70997
ID ADD70997 standard; DNA; 3597 BP.
XX
AC ADD70997;
XX
DT 15-JAN-2004 (first entry)
XX
DE Human attractin gene SEQ ID NO:1.
XX
KW liver cancer; chronic hepatitis; cirrhosis; liver disease; hepatotropic;
KW cytostatic; gene therapy; human; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2003061564-A2.
XX
PD 31-JUL-2003.
XX
PF 20-DEC-2002; 2002WO-US040718.
XX
PR 21-DEC-2001; 2001US-0341815P.
PR 31-DEC-2001; 2001US-0343185P.
XX
PA (GENE-) GENE LOGIC INC.
PA (LGBI-) LG BIOMEDICAL INST.
XX
PI Koh SS, Liu Q, Chung H, Zeng W, Lee B, Yeramilli S, Song SY;
XX
DR WPI; 2003-663343/62.
XX
PT Diagnosing liver cancer cells, useful for treating liver cancer
PT associated with chronic hepatitis or cirrhosis comprises detecting the
PT level of expression in a tissue sample of one or more genes associated
PT with cancerous liver tissues.
XX
PS Claim 1; SEQ ID NO 1; 176pp; English.
XX
CC The present invention describes a method for diagnosing liver cancer
CC cells comprising detecting the level of expression in a tissue sample of
CC one or more genes given in the specification (see ADD70997 to ADD71105),
CC where differential expression of the genes is indicative of liver cancer.
CC Also described: (1) detecting the progression of liver cancer in a
CC patient; (2) monitoring the treatment of a patient with liver cancer; (3)
CC treating a patient with liver cancer; (4) typing a liver disease in a
CC patient; (5) detecting the presence or progression of liver cancer in a
CC patient with chronic hepatitis or cirrhosis; (6) differentiating liver
CC cancer related to chronic hepatitis from liver cancer related to
CC cirrhosis; (7) screening for an agent capable of modulating the onset or
CC progression of liver cancer; (8) a composition comprising at least two
CC oligonucleotides comprising a sequence that specifically hybridises to
CC any of the genes; (9) a solid support comprising the at least two
CC oligonucleotides; (10) a computer system comprising a database containing
CC information identifying the level in liver tissue of a set of genes; (11)
CC a method for using the computer system to present information identifying
CC the expression level in tissue or cell of any of the genes; and (12) a
CC therapeutic agent for slowing or halting the progression of liver cancer.
CC The methods are useful for treating liver cancer associated with chronic
CC hepatitis or cirrhosis. The present sequence represents a specifically
CC claimed human gene sequence which is used in the exemplification of the
CC present invention.
XX
SQ Sequence 3597 BP; 951 A; 822 C; 898 G; 926 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	3597
Score:	874.00	Matches:	1160
Percent Similarity:	99.49%	Conservative:	0
Best Local Similarity:	99.49%	Mismatches:	3
Query Match:	61.16%	Indels:	6
DB:	9	Gaps:	0

US-09-787-097-12 (1-1429) x ADD70997 (1-3597)

```
Qy        105 CysValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTrp 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 TGTGTCAACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGG 150

Qy        125 ValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPhe 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 GTGGGCGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTT 210

Qy        145 ValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeuIleGluGly 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 GTGACAGATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGA 270

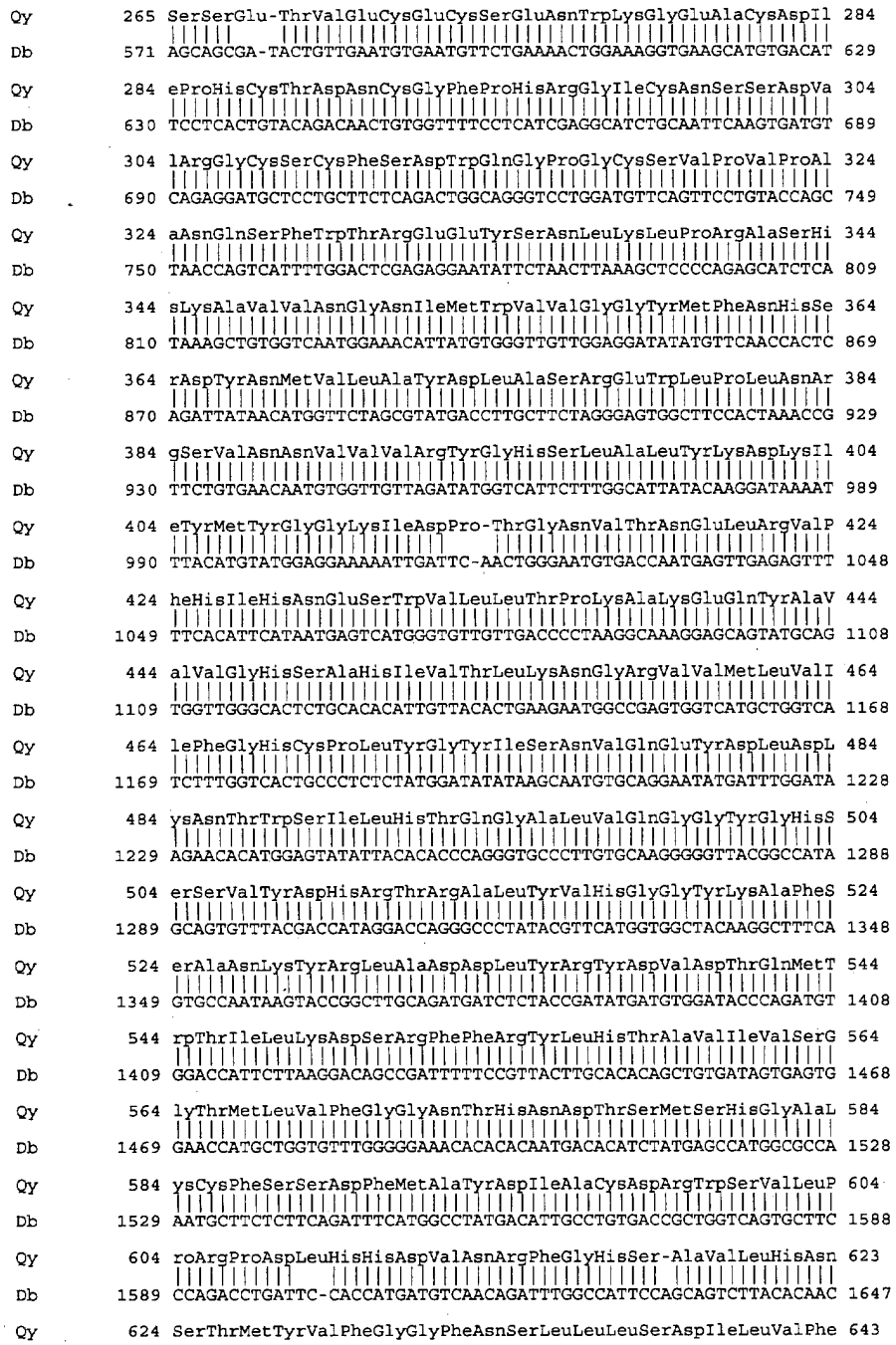
Qy        165 GlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAsp 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 CAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGAC 330

Qy        185 HisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaAlaPheSerGly 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 CATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGC 390

Qy        205 LeuIleValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGly 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 CTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGT 450

Qy        225 TyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 TATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACT 510

Qy        245 TyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsn 264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 TACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAAT 570
```


Qy 265 SerSerGlu-ThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIl 284
Db 571 AGCAGCGA-TACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACAT 629
Qy 284 eProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspVa 304
Db 630 TCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGT 689
Qy 304 lArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAl 324
Db 690 CAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGC 749
Qy 324 aAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHi 344
Db 750 TAACCAGTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCA 809
Qy 344 sLysAlaValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHisSe 364
Db 810 TAAAGCTGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTC 869
Qy 364 rAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnAr 384
Db 870 AGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCG 929
Qy 384 gSerValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIl 404
Db 930 TTCTGTGAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAAT 989
Qy 404 eTyrMetTyrGlyGlyLysIleAspPro-ThrGlyAsnValThrAsnGluLeuArgValP 424
Db 990 TTACATGTATGGAGGAAAAATTGATTC-AACTGGGAATGTGACCAATGAGTTGAGAGTTT 1048
Qy 424 heHisIleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaV 444
Db 1049 TTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAG 1108
Qy 444 alValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValI 464
Db 1109 TGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCA 1168
Qy 464 lePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspL 484
Db 1169 TCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATA 1228
Qy 484 ysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisS 504
Db 1229 AGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATA 1288
Qy 504 erSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheS 524
Db 1289 GCAGTGTTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCA 1348
Qy 524 erAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetT 544
Db 1349 GTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGT 1408
Qy 544 rpThrIleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerG 564
Db 1409 GGACCATTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTG 1468
Qy 564 lyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaL 584
Db 1469 GAACCATGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGCCATGGCGCCA 1528
Qy 584 ysCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuP 604
Db 1529 AATGCTTCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTC 1588
Qy 604 roArgProAspLeuHisHisAspValAsnArgPheGlyHisSer-AlaValLeuHisAsn 623
Db 1589 CCAGACCTGATTC-CACCATGATGTCAACAGATTTGGCCATTCCAGCAGTCTTACACAAC 1647
Qy 624 SerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPhe 643
Db 1648 AGCACCATGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTC 1707
Qy 644 ThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGly 663
Db 1708 ACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGT 1767
Qy 664 IleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAsp 683
Db 1768 ATTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGAT 1827
Qy 684 GluGlnGluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArg 703
Db 1828 GAACAAGAAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGA 1887
Qy 704 CysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCys 723
Db 1888 TGTGACCAGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGC 1947
Qy 724 AsnAspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArg 743
Db 1948 AATGACCATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGG 2007
Qy 744 TyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArg 763
Db 2008 TATGAGAATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGG 2067
Qy 764 SerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeu 783
Db 2068 AGCTGTGCCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTG 2127
Qy 784 ProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThr 803
Db 2128 CCCGAAAATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACT 2187
Qy 804 ThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeu 823
Db 2188 ACTGCCAAGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTG 2247
Qy 824 AlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGln 843
Db 2248 GCTTCTCTTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAG 2307
Qy 844 SerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsn 863
Db 2308 TCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAAT 2367
Qy 864 ValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMet 883
Db 2368 GTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATG 2427
Qy 884 ProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGly 903
Db 2428 CCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGA 2487
Qy 904 LeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsn 923
Db 2488 CTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAAC 2547
Qy 924 HisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThr 943
Db 2548 CACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACC 2607
Qy 944 SerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAla 963
Db 2608 AGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCC 2667
Qy 964 TyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSerThrCysPro 983
Db 2668 TATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCC 2727
Qy 984 ProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGly 1003

```
Db       2728 CCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGC 2787

Qy       1004 TrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGly 1023
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2788 TGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGA 2847

Qy       1024 ProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsn 1043
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2848 CCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAAT 2907

Qy       1044 SerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCys 1063
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2908 TCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGC 2967

Qy       1064 GlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeu 1083
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2968 CAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTG 3027

Qy       1084 ThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGly 1103
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3028 ACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGA 3087

Qy       1104 GlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGly 1123
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3088 GGGAAATGTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGC 3147

Qy       1124 LysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGlu 1143
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3148 AAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAA 3207

Qy       1144 AsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyr 1163
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3208 AATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTAT 3267

Qy       1164 GlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIleAsnPheVal 1183
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3268 CAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTG 3327

Qy       1184 AlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPhe 1203
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3328 GCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTC 3387

Qy       1204 AsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMet 1223
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3388 AACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATG 3447

Qy       1224 ProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPhe 1243
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3448 CCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTT 3507

Qy       1244 AspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpPro 1263
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3508 GATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCC 3567

Qy       1264 IleLysIleGln 1267
              |||||||||||
Db       3568 ATCAAAATTCAG 3579
```

RESULT 10
AAZ91922
ID AAZ91922 standard; cDNA; 2625 BP.
XX
AC AAZ91922;
XX
DT 08-JUN-2000 (first entry)
XX
DE Human mahogany protein coding sequence #4.
XX
KW Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;
KW weight regulation; cell therapy; body weight disorder; cachexia;
KW anorexia; hyperpigmentation; increased metabolic rate disorder;
KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.
XX
OS Homo sapiens.
XX
PN WO200005373-A2.
XX
PD 03-FEB-2000.
XX
PF 21-JUL-1999; 99WO-US016484.
XX
PR 21-JUL-1998; 98US-0093630P.
PR 20-OCT-1998; 98US-0104978P.
PR 05-FEB-1999; 99US-00245041.
XX
PA (MILL-) MILLENIUM PHARM INC.
XX
PI Moore K, Nagle DL;
XX
DR WPI; 2000-195103/17.
DR P-PSDB; AAY81809.
XX
PT New human and murine mahogany genes, useful, e.g. for diagnosis and
PT treatment of body weight disorders.
XX
PS Claim 1; Fig 20a; 188pp; English.
XX
CC This sequence represents a human mahogany gene of the invention. The
CC mahogany genes are used: (i) to produce recombinant mahogany (mg)
CC proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming
CC therapeutics; (iii) as a source of diagnostic probes and primers for
CC detecting expression of mg genes or mutations, regulatory defects, in
CC this gene, or for isolation of related sequences; and (iv) in (cell-
CC based) gene therapy. (II) are used to raise specific antibodies (Ab); to
CC identify other (extra)cellular products involved in weight regulation,
CC and to screen for agents that disrupt interaction between (II) and other
CC macromolecules. The Ab are used to detect abnormal levels (or function)
CC of (II) (for diagnosis, prognosis or monitoring of treatment); to
CC evaluate (II)-expressing cells intended for cell therapy, and as
CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the
CC mg polypeptide) are used to identify agents (A) that modulate mg
CC activity. (A) are potentially useful for the treatment of body weight
CC disorders, particularly obesity, cachexia or anorexia, or other
CC conditions associated with the mg gene such as hyperpigmentation,
CC hyperphagia and disorders that result in increased metabolic rate
XX
SQ Sequence 2625 BP; 660 A; 605 C; 694 G; 666 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	2625
Score:	496.00	Matches:	696
Percent Similarity:	99.43%	Conservative:	0
Best Local Similarity:	99.43%	Mismatches:	2
Query Match:	34.71%	Indels:	4
DB:	3	Gaps:	0

US-09-787-097-12 (1-1429) x AAZ91922 (1-2625)

```
Qy         87 AlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGluCysAspArgProCysVal 106
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 GCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAAGGAATGTGACCGGCCGTGTGTC 315

Qy        107 AsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTrpValGly 126
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        316 AACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGC 375

Qy        127 GluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPheValThr 146
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        376 GAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACA 435

Qy        147 AspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeuIleGluGlyGlnPro 166
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        436 GATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCA 495

Qy        167 AsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLeu 186
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        496 AATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCATTTA 555

Qy        187 TyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaAlaPheSerGlyLeuIle 206

Db        556 TATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATT 615

Qy        207 ValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGlyTyrAla 226

Db        616 GTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCC 675

Qy        227 LeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThrTyrSer 246

Db        676 TTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTACAGT 735

Qy        247 PheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSerSer 266

Db        736 TTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGC 795

Qy        267 Glu-ThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIleProHi 286

Db        796 GA-TACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCA 854

Qy        286 sCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspValArgGl 306

Db        855 CTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGG 914

Qy        306 yCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAlaAsnGl 326

Db        915 ATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCA 974

Qy        326 nSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHisLysAl 346

Db        975 GTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGC 1034

Qy        346 aValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHisSerAspTy 366

Db       1035 TGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTA 1094

Qy        366 rAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSerVa 386

Db       1095 TAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGT 1154

Qy        386 lAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyrMe 406

Db       1155 GAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACAT 1214

Qy        406 tTyrGlyGlyLysIleAspPro-ThrGlyAsnValThrAsnGluLeuArgValPheHisI 426

Db       1215 GTATGGAGGAAAAATTGATTC-AACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACA 1273

Qy        426 leHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaValValG 446

Db       1274 TTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTG 1333

Qy        446 lyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValIlePheG 466

Db       1334 GGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTG 1393

Qy        466 lyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsnT 486

Db       1394 GTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACA 1453

Qy        486 hrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisSerServ 506

Db       1454 CATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTG 1513

Qy        506 alTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaA 526

Db       1514 TTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCCA 1573

Qy        526 snLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetTrpThrI 546

Db       1574 ATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCA 1633

Qy        546 leLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerGlyThrM 566

Db       1634 TTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACCA 1693

Qy        566 etLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysP 586

Db       1694 TGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATGCT 1753

Qy        586 heSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuProArgP 606

Db       1754 TCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGAC 1813

Qy        606 roAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrM 626

Db       1814 CTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACAACAGCACCA 1873

Qy        626 etTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSerG 646

Db       1874 TGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCGG 1933

Qy        646 luGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgC 666

Db       1934 AACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGT 1993

Qy        666 ysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnG 686

Db       1994 GTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAAG 2053

Qy        686 luGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArgCysAspG 706

Db       2054 AAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACC 2113

Qy        706 lnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAspH 726

Db       2114 AGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGCAATGACC 2173

Qy        726 isCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGluA 746

Db       2174 ATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGA 2233

Qy        746 snCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArgSerCysA 766

Db       2234 ATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTG 2293

Qy        766 laLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuPro 784

Db       2294 CCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCC 2349
```

RESULT 11
AAZ91919
ID AAZ91919 standard; cDNA; 6373 BP.
XX
AC AAZ91919;
XX
DT 08-JUN-2000 (first entry)
XX
DE Human mahogany protein coding sequence #1.
XX
KW Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;
KW weight regulation; cell therapy; body weight disorder; cachexia;
KW anorexia; hyperpigmentation; increased metabolic rate disorder;
KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT misc_difference 448
FT /*tag= a
FT /note= "represented in specification as :"
FT misc_difference 615
FT /*tag= b
FT /note= "represented in specification as :"

```
FT   misc_difference 683
FT                   /*tag=  c
FT                   /note= "represented in specification as :"
XX
PN   WO200005373-A2.
XX
PD   03-FEB-2000.
XX
PF   21-JUL-1999;   99WO-US016484.
XX
PR   21-JUL-1998;   98US-0093630P.
PR   20-OCT-1998;   98US-0104978P.
PR   05-FEB-1999;   99US-00245041.
XX
PA   (MILL-) MILLENIUM PHARM INC.
XX
PI   Moore K,  Nagle DL;
XX
DR   WPI; 2000-195103/17.
DR   P-PSDB; AAY81806.
XX
PT   New human and murine mahogany genes, useful, e.g. for diagnosis and
PT   treatment of body weight disorders.
XX
PS   Claim 1; Fig 10a; 188pp; English.
XX
CC   This sequence represents a human mahogany gene of the invention. The
CC   mahogany genes are used: (i) to produce recombinant mahogany (mg)
CC   proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming
CC   therapeutics; (iii) as a source of diagnostic probes and primers for
CC   detecting expression of mg genes or mutations, regulatory defects, in
CC   this gene, or for isolation of related sequences; and (iv) in (cell-
CC   based) gene therapy. (II) are used to raise specific antibodies (Ab); to
CC   identify other (extra)cellular products involved in weight regulation,
CC   and to screen for agents that disrupt interaction between (II) and other
CC   macromolecules. The Ab are used to detect abnormal levels (or function)
CC   of (II) (for diagnosis, prognosis or monitoring of treatment); to
CC   evaluate (II)-expressing cells intended for cell therapy, and as
CC   therapeutic mg inhibitors. Cells that express the mg gene (or contain the
CC   mg polypeptide) are used to identify agents (A) that modulate mg
CC   activity. (A) are potentially useful for the treatment of body weight
CC   disorders, particularly obesity, cachexia or anorexia, or other
CC   conditions associated with the mg gene such as hyperpigmentation,
CC   hyperphagia and disorders that result in increased metabolic rate
XX
SQ   Sequence 6373 BP; 1736 A; 1468 C; 1460 G; 1699 T; 0 U; 10 Other;

Alignment Scores:
Pred. No.:              0              Length:          6373
Score:                  467.00         Matches:         480
Percent Similarity:     99.79%         Conservative:    0
Best Local Similarity:  99.79%         Mismatches:      0
Query Match:            32.68%         Indels:          1
DB:                     3              Gaps:            0

US-09-787-097-12 (1-1429) x AAZ91919 (1-6373)

Qy        950 MetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrValAlaSerPhePro 969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 ATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCT 676

Qy        970 PheGly-GlnCysMetGluTrpTyrThrMetSerThrCysProProGluAsnCysSerGl 989
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 TTTGGNCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGG 736

Qy        989 yTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCysThrAspProSe 1009
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 CTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAG 796

Qy       1009 rAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyProValLysMetProSe 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 CAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTC 856

Qy       1029 rGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerSerMetCysLeuGl 1049
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 GCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGA 916

Qy       1049 uAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnCysAsnGlyHisSe 1069
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 GGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAG 976

Qy       1069 rLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrThrGlyLysHisCy 1089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 TAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTG 1036

Qy       1089 sGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyGlyLysCysGlnProCy 1109
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1037 CGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATG 1096

Qy       1109 sLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysPheCysThrTh 1129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1097 CAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCAC 1156

Qy       1129 rLysGlyValLysGlyAspGluCysGlnLeuCysGluValGluAsnArgTyrGlnGlyAs 1149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1157 CAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAA 1216

Qy       1149 nProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrGlnPheThrPheSerLe 1169
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1217 CCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCT 1276

Qy       1169 uSerGlnGluAspAspArgTyrTyrThrAlaIleAsnPheValAlaThrProAspGluGl 1189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1277 ATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACA 1336

Qy       1189 nAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTr 1209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1337 AAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTG 1396

Qy       1209 pAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProValValSerLysTh 1229
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1397 GGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAAC 1456

Qy       1229 rAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPheAspPheArgAsnHisPr 1249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1457 CAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCC 1516

Qy       1249 oAsnIleThrPhePheValTyrValSerAsnPheThrTrpProIleLysIleGlnIleAl 1269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1517 AAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGATTGC 1576

Qy       1269 aPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePheValThrPhePheSerCy 1289
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1577 CTTCTCTCAGCACAGCAATTTTATGGACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTG 1636

Qy       1289 sPheLeuSerLeuLeuLeuValAlaAlaValValTrpLysIleLysGlnSerCysTrpAl 1309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1637 TTTCCTCTCTTTGCTCCTGGTGGCTGCTGTGGTTTGGAAGATCAAACAAAGTTGTTGGGC 1696

Qy       1309 aSerArgArgArgGluGlnLeuLeuArgGluMetGlnGlnMetAlaSerArgProPheAl 1329
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1697 CTCCAGACGTAGAGAGCAACTTCTTCGAGAGATGCAACAGATGGCCAGCCGTCCCTTTGC 1756

Qy       1329 aSerValAsnValAlaLeuGluThrAspGluGluProProAspLeuIleGlyGlySerIl 1349
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1757 CTCTGTAAATGTCGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTATTGGGGGGAGTAT 1816

Qy       1349 eLysThrValProLysProIleAlaLeuGluProCysPheGlyAsnLysAlaAlaValLe 1369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1817 AAAGACTGTTCCCAAACCCATTGCACTGGAGCCGTGTTTTGGCAACAAAGCCGCTGTCCT 1876

Qy       1369 uSerValPheValArgLeuProArgGlyLeuGlyGlyIleProProProGlyGlnSerGl 1389
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1877 CTCTGTGTTTGTGAGGCTCCCTCGAGGCCTGGGTGGCATCCCTCCTCCTGGGCAGTCAGG 1936

Qy       1389 yLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetProIleValTyrLysGl 1409
```

```
Db       1937 TCTTGCTGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGATAGTGTACAAGGA 1996

Qy       1409 uLysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnProGlyThrCysIl 1429

Db       1997 GAAGTCAGGAGCCGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGCCTGGGACCTGCAT 2056

Qy       1429 e 1429

Db       2057 C 2057
```

RESULT 12
AAS72660
ID AAS72660 standard; cDNA; 3490 BP.
XX
AC AAS72660;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #8464.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG08473.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 8464; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 3490 BP; 931 A; 772 C; 858 G; 929 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3.2e-245	Length:	3490
Score:	289.00	Matches:	528
Percent Similarity:	98.69%	Conservative:	0
Best Local Similarity:	98.69%	Mismatches:	2
Query Match:	20.22%	Indels:	7
DB:	5	Gaps:	0

US-09-787-097-12 (1-1429) x AAS72660 (1-3490)

```
Qy        137 ArgLeuThrGlySerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThr 156

Db         22 AGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAAATACAAAACG 81

Qy        157 LysCysThrTrpLeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHis 176

Db         82 AAGTGCACGTGGCTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCAT 141

Qy        177 PheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAla 196

Db        142 TTTGCTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGGACTCAATTTATGCA 201

Qy        197 ProLeuValAlaAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrVal 216

Db        202 CCGCTAGTTGCTGCATTTAGTGGCCTCATTGTTCCTGAGAGAGATGGCAATGAGACTGTC 261

Qy        217 ProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyr 236

Db        262 CCTGAGGTTGTTGCCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTAT 321

Qy        237 AsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGly 256

Db        322 AATTTGACTGGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGC 381

Qy        257 ArgGlyGluCysLysIleSerAsnSerSerGlu-ThrValGluCysGluCysSerGluAs 276

Db        382 CGAGGAGAGTGTAAGATCAGTAATAGCAGCGA-TACTGTTGAATGTGAATGTTCTGAAAA 440

Qy        276 nTrpLysGlyGluAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisAr 296

Db        441 CTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCG 500

Qy        296 gGlyIleCysAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyPr 316

Db        501 AGGCATCTGCAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCC 560

Qy        316 oGlyCysSerValProValProAlaAsnGlnSerPheTrp-ThrArgGluGluTyrSerA 336

Db        561 TGGATGTTCAGTTCCTGTACCAGCTAACCAGTCATTTTGGGACTCGAGAGGAATATTCTA 620

Qy        336 sn-LeuLysLeu-ProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpVa 355

Db        621 ACTTTAAAGCTCCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTATGTGGGT 680

Qy        355 lValGlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAl 375

Db        681 TGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGC 740

Qy        375 aSerArgGluTrpLeuProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHi 395

Db        741 TTCTAGGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTTAGATATGGTCA 800

Qy        395 sSerLeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLysIleAspPro-ThrG 415

Db        801 TTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTC-AACTG 859

Qy        415 lyAsnValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuT 435

Db        860 GGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGA 919

Qy        435 hrProLysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuL 455

Db        920 CCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGA 979
```

```
Qy        455 ysAsnGlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleS 475
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        980 AGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAA 1039

Qy        475 erAsnValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyA 495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1040 GCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTG 1099

Qy        495 laLeuValGlnGlyGlyTyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuT 515
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1100 CCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAGGGCCCTAT 1159

Qy        515 yrValHisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuT 535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1160 ACGTTCATGGTGGCTACAAGGCTTTCAGTGCCAATAAGTACCGGCTTGCAGATGATCTCT 1219

Qy        535 yrArgTyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgT 555
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1220 ACCGATATGATGTGGATACCCAGATGTGGACCATTCTTAAGGACAGCCGATTTTTCCGTT 1279

Qy        555 yrLeuHisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisA 575
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1280 ACTTGCACACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGAGGAAACACACACA 1339

Qy        575 snAspThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspI 595
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1340 ATGACACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTTCATGGCCTATGACA 1399

Qy        595 leAlaCysAspArgTrpSerValLeuProArgProAspLeuHisHisAspValAsnArgP 615
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1400 TTGCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGAT 1459

Qy        615 heGlyHisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerL 635
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1460 TTGGCCATTCAGCAGTCTTACACAACAGCACCATGTATGTGTTCGGTGGTTTCAATAGTC 1519

Qy        635 euLeuLeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluA 655
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1520 TCCTCCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAG 1579

Qy        655 laAlaCysLeuAlaAlaGlyProGlyIleArgCys 666
              ||||||||||||||||||||||||||||||||||
Db       1580 CCGCTTGTTTAGCAGCAGGACCTGGTATTCGGTGT 1614
```

```
RESULT 13
ADB53249
ID   ADB53249 standard; DNA; 8739 BP.
XX
AC   ADB53249;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3791.
XX
KW   toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KW   toxicity marker; toxicity progression; drug screening;
KW   primary rat hepatocyte toxicity modelling; gene; ds.
XX
OS   Rattus norvegicus.
XX
PN   WO2003065993-A2.
XX
PD   14-AUG-2003.
XX
PF   04-FEB-2003; 2003WO-US003482.
XX
PR   04-FEB-2002; 2002US-0353171P.
PR   13-MAR-2002; 2002US-0363534P.
PR   08-APR-2002; 2002US-0370248P.
PR   10-APR-2002; 2002US-0371134P.
PR   10-APR-2002; 2002US-0371135P.
PR   10-APR-2002; 2002US-0371150P.
PR   11-APR-2002; 2002US-0371413P.
PR   19-APR-2002; 2002US-0373601P.
PR   19-APR-2002; 2002US-0373602P.
PR   22-APR-2002; 2002US-0374139P.
PR   08-MAY-2002; 2002US-0378370P.
PR   09-MAY-2002; 2002US-0378652P.
PR   09-MAY-2002; 2002US-0378653P.
PR   09-MAY-2002; 2002US-0378665P.
PR   09-JUL-2002; 2002US-0394230P.
PR   09-JUL-2002; 2002US-0394253P.
PR   04-SEP-2002; 2002US-0407688P.
PR   28-JAN-2003; 2003US-0442900P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Mendrick D,  Porter M,  Johnson K,  Higgs B,  Castle A,  Orr M;
PI   Elashoff M;
XX
DR   WPI; 2003-731472/69.
XX
PT   Determining if a compound induces a toxic effect on a tissue or cell, for
PT   identifying hepatotoxic compounds, comprises comparing a gene expression
PT   profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT   mean values.
XX
PS   Claim 44; SEQ ID NO 3791; 874pp; English.
XX
CC   The present invention describes a method for determining whether a
CC   compound induces a toxic effect on a tissue or cell. The method comprises
CC   preparing a gene expression profile of a tissue or cell sample exposed to
CC   the compound, and comparing the gene expression profile to a database
CC   comprising data or information on the Tox mean and non-Tox mean value.
CC   The method is useful for predicting or identifying at least one toxic
CC   effect, particularly hepatotoxicity, of a test or unknown compound. The
CC   genes listed in the specification are useful as diagnostic or toxicity
CC   markers for the prediction or identification of the physiological state
CC   of tissue or cell sample that has been exposed to a compound, or to
CC   identify or predict the toxic effects of a compound or an agent. These
CC   may also be used as markers for monitoring toxicity progression or for
CC   drug screening. The present sequence represents a primary rat hepatocyte
CC   toxicity modelling related gene sequence from the present invention.
XX
SQ   Sequence 8739 BP; 2288 A; 2115 C; 2109 G; 2227 T; 0 U; 0 Other;
```

Alignment Scores:			
Pred. No.:	2.09e-171	Length:	8739
Score:	206.00	Matches:	568
Percent Similarity:	95.62%	Conservative:	0
Best Local Similarity:	95.62%	Mismatches:	13
Query Match:	14.42%	Indels:	26
DB:	9	Gaps:	0

US-09-787-097-12 (1-1429) x ADB53249 (1-8739)

```
Qy        849 SerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2614 TCCAAGCTCACTCTGACTCCATGGGTTGGTCTTCGGAAGATCAATGTATCTTACTGGTGC 2673

Qy        869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2674 TGGGAGGATATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGT 2733

Qy        889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2734 GATGCCGGCTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGACTGAAGGCTGCAACC 2793

Qy        909 CysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSerAlaLysGln 928
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2794 TGCATCAACCCACTCAATGGAAGTGTTTGTGAAAGGCCTGCAAACCACAGTGCCAAGCAG 2853

Qy        929 CysArgThrProCysAlaLeuArgThrAlaCysGlyAsp-CysThrSerGly-SerSerG 948
              ||||||||||||||||||||||||||||||||||||    ||||||||    |||||||
Db       2854 TGCCGGACACCCTGTGCCCTGCGGACAGCATGTGGTGA-GTGCACCAGCAG-CAGCTCCG 2911
```

```
Qy       948 luCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrValAlaSerP 968
Db      2912 AATGCATGTGGTGCAGCAACATGAAGCAGTGCGTAGACTCCAATGCCTACGTAGCCTCCT 2971

Qy       968 heProPheGlyGlnCysMetGluTrpTyrThrMetSerThr-CysProProGluAsnCys 987
Db      2972 TTCCTTTCGGCCAGTGTATGGAGTGGTATACAATGAGCAG-CTGCCCACCTGAAAATTGC 3030

Qy       988 SerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCysThrAsp 1007
Db      3031 TCTGGCTACTGTACCTGCAGCCATTGCTTGGAGCAGCCAGGCTGTGGTTGGTGTACTGAT 3090

Qy      1008 ProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyProValLysMet 1027
Db      3091 CCTAGCAATACTGGGAAAGGCAAATGTATTGAGGGAAGCTATAAAGGACCCGTGAAGATG 3150

Qy      1028 ProSerGlnAlaPro--ThrGlyAsnPhe-TyrProGlnProLeuLeuAsnSerSerMet 1046
Db      3151 CCTTCACA-TG-CCTCTACAGGAAATGT-GTACCCACAGCCCCTTCTGAACTCCAGCATG 3207

Qy      1047 CysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnCysAsn 1066
Db      3208 TGTCTAGAGGACAGCAGATACAACTGGTCTTTTATTCACTGTCCAGCTTGCCAGTGCAAT 3267

Qy      1067 GlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsn-LeuThrThrGl 1086
Db      3268 GGACACAGCAAGTGCATCAACCAGAGCATCTGTGAGAAGTGTGAGGA-CCTGACCACAGG 3326

Qy      1086 yLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyGlyLysCy 1106
Db      3327 CAAGCACTGCGAGACCTGCATATCTGGCTTCTATGGTGACCCGACTAATGGAGGCAAATG 3386

Qy      1106 sGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysPh 1126
Db      3387 TCAGCCATGCAAGTGCAACGGGCACGCATCACTGTGCAACACCAACACTGGCAAGTGCTT 3446

Qy      1126 eCysThrThrLysGlyValLysGlyAsp-GluCysGlnLeuCysGluValGluAsnArgT 1146
Db      3447 CTGTACCACCAAGGGCGTCAAGGGAGA-GGAGTGCCAGCTATGTGAGGTAGAAAATCGAT 3505

Qy      1146 yrGlnGlyAsnProLeuArg-GlyThrCysTyrTyrThrLeuLeuIleAspTyrGlnPhe 1165
Db      3506 ACCAGGGAAACCCTCTCAA-AGGAACATGTTACTATACCCTTCTCATTGACTATCAGTTC 3564

Qy      1166 ThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIleAsnPheValAlaThr 1185
Db      3565 ACTTTTAGCCTGTCCCAGGAAGACGACCGCTATTACACAGCCATCAATTTTGTGGCTACT 3624

Qy      1186 ProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeu 1205
Db      3625 CCTGATGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAAAACTTCAACCTC 3684

Qy      1206 AsnIleThrTrpAlaAla-SerPheSer-AlaGlyThrGlnAla-GlyGluGluMet-Pr 1224
Db      3685 AACATCACCTGGGCCAC-CAGTTTCCC-AGCTGGAACCCAGAC-TGGAGAAGAGGT-GCC 3740

Qy      1224 oValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPheAs 1244
Db      3741 TGTTGTTTCAAAAACCAACATCAAGGAATACAAAGATAGCTTCTCTAATGAGAAATTTGA 3800

Qy      1244 pPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpProIl 1264
Db      3801 TTTTCGCAACCATCCAAACATTACTTTCTTTGTTTATGTCAGTAATTTCACTTGGCCCAT 3860

Qy      1264 eLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePheVa 1284
Db      3861 CAAAATTCAAATTGCCTTCTCCCAGCACAGCAATTTTATGGACCTGGTGCAGTTCTTCGT 3920

Qy      1284 lThrPhePheSerCysPheLeuSerLeuLeuLeuValAlaAlaValValTrpLysIleLy 1304
Db      3921 GACTTTCTTCAGCTGTTTCCTCTCTCTGCTTCTGGTGGCGGCAGTGGTCTGGAAGATCAA 3980

Qy      1304 sGlnSerCysTrpAlaSerArgArgArgGluGlnLeuLeuArgGluMetGlnGlnMetAl 1324
Db      3981 GCAGAGCTGTTGGGCATCCAGGAGGAGAGAGCAACTTCTTCGGGAGATGCAACAGATGGC 4040

Qy      1324 aSerArgProPheAlaSerValAsnValAlaLeuGluThrAspGluGluProProAspLe 1344
Db      4041 CAGCCGCCCTTTTGCTTCTGTAAATGTTGCCTTGGAAACAGATGAAGAACCTCCTGATCT 4100

Qy      1344 uIleGlyGlySerIleLysThrValProLysProIleAlaLeuGluProCysPheGlyAs 1364
Db      4101 CATTGGGGGAAGTATAAAGACTGTTCCTAAGCCCATTGCCCTGGAGCCCTGCTTTGGTAA 4160

Qy      1364 nLysAlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyGlyIleProPr 1384
Db      4161 CAAAGCTGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGACTGGGAGGAATCCCTCC 4220

Qy      1384 oProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetPr 1404
Db      4221 TCCTGGTCAGTCAGGTCTTGCTGTGGCCAGTGCCCTGGTGGACATTTCTCAGCAGATGCC 4280

Qy      1404 oIleValTyrLysGluLysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaGl 1424
Db      4281 AATAGTGTACAAGGAGAAGTCAGGAGCTGTAAGGAACCGGAAGCAGCAGCCCCCTGCACA 4340

Qy      1424 nProGlyThrCysIle 1429
Db      4341 GCCTGGAACCTGCATT 4356
```

RESULT 14
AAZ91916
ID AAZ91916 standard; cDNA; 8827 BP.
XX
AC AAZ91916;
XX
DT 08-JUN-2000 (first entry)
XX
DE Wild type (C57BL/6J) murine mahogany protein coding sequence.
XX
KW Mahogany gene; mouse; mg gene; regulatory defect; gene therapy; obesity;
KW weight regulation; cell therapy; body weight disorder; cachexia;
KW anorexia; hyperpigmentation; increased metabolic rate disorder;
KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.
XX
OS Mus sp.
XX
PN WO200005373-A2.
XX
PD 03-FEB-2000.
XX
PF 21-JUL-1999; 99WO-US016484.
XX
PR 21-JUL-1998; 98US-0093630P.
PR 20-OCT-1998; 98US-0104978P.
PR 05-FEB-1999; 99US-00245041.
XX
PA (MILL-) MILLENIUM PHARM INC.
XX
PI Moore K, Nagle DL;
XX
DR WPI; 2000-195103/17.
DR P-PSDB; AAY81803.
XX
PT New human and murine mahogany genes, useful, e.g. for diagnosis and
PT treatment of body weight disorders.
XX
PS Claim 1; Fig 2a; 188pp; English.
XX
CC This sequence represents a murine mahogany gene of the invention. The
CC mahogany genes are used: (i) to produce recombinant mahogany (mg)
CC proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming
CC therapeutics; (iii) as a source of diagnostic probes and primers for
CC detecting expression of mg genes or mutations, regulatory defects, in
CC this gene, or for isolation of related sequences; and (iv) in (cell-

```
CC   based) gene therapy. (II) are used to raise specific antibodies (Ab); to
CC   identify other (extra)cellular products involved in weight regulation,
CC   and to screen for agents that disrupt interaction between (II) and other
CC   macromolecules. The Ab are used to detect abnormal levels (or function)
CC   of (II) (for diagnosis, prognosis or monitoring of treatment); to
CC   evaluate (II)-expressing cells intended for cell therapy, and as
CC   therapeutic mg inhibitors. Cells that express the mg gene (or contain the
CC   mg polypeptide) are used to identify agents (A) that modulate mg
CC   activity. (A) are potentially useful for the treatment of body weight
CC   disorders, particularly obesity, cachexia or anorexia, or other
CC   conditions associated with the mg gene such as hyperpigmentation,
CC   hyperphagia and disorders that result in increased metabolic rate
XX
SQ   Sequence 8827 BP; 2304 A; 2077 C; 2162 G; 2284 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                2.79e-149      Length:          8827
Score:                    181.00         Matches:         575
Percent Similarity:       95.67%         Conservative:    0
Best Local Similarity:    95.67%         Mismatches:      13
Query Match:              12.67%         Indels:          26
DB:                       3              Gaps:            0

US-09-787-097-12 (1-1429) x AAZ91916 (1-8827)

Qy        842 MetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLys 861
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2790 ATGCAATCATCTCAAAGTATGTCCAAGCTCACTCTGACTCCATGGGTTGGTCTTCGGAAG 2849

Qy        862 IleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGln 881
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2850 ATCAATGTGTCTTACTGGTGCTGGGAGGATATGTCTCCATTCACAAATAGTTTGCTGCAG 2909

Qy        882 TrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThr 901
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2910 TGGATGCCATCTGAGCCCAGTGATGCTGGCTTCTGTGGGATCTTGTCAGAGCCTAGTACT 2969

Qy        902 ArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro 921
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2970 CGGGGATTAAAGGCTGCAACCTGCATCAACCCTCTCAATGGCAGCGTCTGTGAAAGGCCT 3029

Qy        922 AlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAsp 941
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3030 GCAAACCACAGTGCCAAGCAGTGCCGGACACCATGTGCCCTGCGGACAGCGTGTGGCGA- 3088

Qy        942 -CysThrSerGly-SerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspS 961
               |||||||||    ||||||||||||||||||||||||||||||||||||||||||||||
Db       3089 GTGCACTAGCAG-CAGCTCGGAGTGCATGTGGTGCAGTAACATGAAGCAGTGTGTGGACT 3147

Qy        961 erAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSerT 981
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3148 CCAATGCCTACGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCA 3207

Qy        981 hr-CysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPro 1000
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3208 G-CTGCCCACCTGAAAATTGCTCTGGCTACTGTACCTGCAGCCATTGCTTGGAGCAGCCA 3266

Qy       1001 GlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySer 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3267 GGCTGTGGTTGGTGTACTGATCCTAGCAATACTGGGAAAGGAAAATGTATTGAGGGCAGC 3326

Qy       1021 TyrLysGlyProValLysMetProSerGlnAlaProThr--GlyAsnPhe-TyrProGln 1039
              |||||||||||||||||||||||||||||||||||||||  ||||||  |||||||||||
Db       3327 TATAAAGGACCTGTGAAGATGCCGTCACAGGCCTC-TG-CAGGAAATGT-GTATCCACAG 3383

Qy       1040 ProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHis 1059
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3384 CCCCTTCTGAACTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCAC 3443

Qy       1060 CysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLys 1079
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3444 TGTCCAGCTTGCCAGTGCAACGGACACAGCAAATGCATCAACCAGAGTATCTGTGAGAAG 3503

Qy       1080 CysGluAsn-LeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAs 1099
              ||||||    ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3504 TGTGAGGA-CCTGACCACGGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTATGGTGA 3562

Qy       1099 pProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAs 1119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3563 CCCGACTAATGGAGGCAAATGTCAGCCATGCAAGTGCAATGGGCACGCATCACTGTGCAA 3622

Qy       1119 nThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLe 1139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3623 CACCAACACCGGCAAGTGCTTCTGTACCACCAAAGGTGTCAAGGGGGACGAGTGCCAGCT 3682

Qy       1139 uCysGluValGluAsnArgTyrGlnGlyAsnProLeuArg-GlyThrCysTyrTyrThrL 1159
              ||||||||||||||||||||||||||||||||||||||    ||||||||||||||||||||
Db       3683 ATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAA-AGGAACATGCTACTATACCC 3741

Qy       1159 euLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrA 1179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3742 TTCTCATTGACTATCAGTTCACCTTTAGCCTGTCCCAGGAAGACGACCGCTACTACACAG 3801

Qy       1179 laIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnA 1199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3802 CCATCAACTTTGTGGCTACTCCTGATGAACAAAACAGGGATTTGGACATGTTCATCAATG 3861

Qy       1199 laSerLysAsnPheAsnLeuAsnIleThrTrpAlaAla-SerPheSer-AlaGlyThrGl 1218
              ||||||||||||||||||||||||||||||||||||    ||||||    ||||||||||
Db       3862 CCTCCAAAAACTTCAACCTCAACATCACCTGGGCCAC-CAGCTTCCC-AGCCGGAACCCA 3919

Qy       1218 nAla-GlyGluGluMet-ProValValSerLysThrAsnIleLysGluTyrLysAspSer 1237
              |   |||||||||    ||||||||||||||||||||||||||||||||||||||||||||
Db       3920 GAC-TGGAGAAGAGGT-GCCTGTTGTTTCAAAAACCAACATCAAGGAATACAAAGATAGC 3977

Qy       1238 PheSerAsnGluLysPheAspPheArgAsnHis-ProAsnIleThrPhePheValTyrVa 1257
              ||||||||||||||||||||||||||||||||    ||||||||||||||||||||||||||
Db       3978 TTCTCTAATGAGAAATTTGATTTTCGCAACGA-TCCAAACATCACTTTCTTTGTTTATGT 4036

Qy       1257 lSerAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMe 1277
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4037 CAGTAATTTCACTTGGCCCATCAAAATTCAGATTGCCTTCTCCCAGCACAGCAACTTCAT 4096

Qy       1277 tAspLeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuLeuValAl 1297
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4097 GGACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCGCTGCTTCTGGTGGC 4156

Qy       1297 aAlaValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeuLe 1317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4157 TGCAGTGGTCTGGAAGATCAAGCAGAGCTGTTGGGCATCCAGGCGGAGAGAGCAACTTCT 4216

Qy       1317 uArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluTh 1337
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4217 TCGGGAGATGCAACAGATGGCCAGCCGCCCCTTTGCTTCTGTAAACGTTGCCTTGGAAAC 4276

Qy       1337 rAspGluGluProProAspLeuIleGlyGlySerIleLysThrValProLysProIleAl 1357
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4277 AGATGAGGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGC 4336

Qy       1357 aLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProAr 1377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4337 ACTGGAGCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCG 4396

Qy       1377 gGlyLeuGlyGlyIleProProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuVa 1397
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4397 AGGCCTGGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGT 4456

Qy       1397 lAspIleSerGlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAsnAr 1417
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4457 GGACATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCG 4516

Qy       1417 gLysGlnGlnProProAlaGlnProGlyThrCysIle 1429
              |||||||||||||||||||||||||||||||||||||
Db       4517 GAAGCAGCAGCCCCCTGCACAGCCTGGGACCTGCATC 4553

RESULT 15
```

```
AAS78781
ID   AAS78781 standard; cDNA; 3165 BP.
XX
AC   AAS78781;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #14585.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG14594.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 14585; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 3165 BP; 1042 A; 705 C; 687 G; 731 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              3.49e-115      Length:          3165
Score:                  142.00         Matches:         142
Percent Similarity:     100.00%        Conservative:    0
Best Local Similarity:  100.00%        Mismatches:      0
Query Match:            9.94%          Indels:          0
DB:                     5              Gaps:            0

US-09-787-097-12 (1-1429) x AAS78781 (1-3165)

Qy       920 ArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCys 939
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db      1219 AGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGT 1278

Qy       940 GlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysVal 959
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db      1279 GGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTG 1338

Qy       960 AspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMet 979
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db      1339 GACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATG 1398

Qy       980 SerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGln 999
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db      1399 AGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAA 1458

Qy      1000 ProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGly 1019
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db      1459 CCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGT 1518

Qy      1020 SerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGln 1039
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db      1519 TCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAG 1578

Qy      1040 ProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHis 1059
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db      1579 CCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCAC 1638

Qy      1060 CysPro 1061
             ||||||
Db      1639 TGTCCA 1644

RESULT 16
AAS76294
ID   AAS76294 standard; cDNA; 3313 BP.
XX
AC   AAS76294;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #12098.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG12107.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 12098; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   ahd in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
```

```
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 3313 BP; 1090 A; 731 C; 724 G; 768 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              3.64e-115      Length:          3313
Score:                  142.00         Matches:         142
Percent Similarity:     100.00%        Conservative:    0
Best Local Similarity:  100.00%        Mismatches:      0
Query Match:            9.94%          Indels:          0
DB:                     5              Gaps:            0

US-09-787-097-12 (1-1429) x AAS76294 (1-3313)

Qy        920 ArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCys 939
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1219 AGGCCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGT 1278

Qy        940 GlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysVal 959
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1279 GGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTG 1338

Qy        960 AspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMet 979
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1339 GACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATG 1398

Qy        980 SerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGln 999
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1399 AGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAA 1458

Qy       1000 ProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGly 1019
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1459 CCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGT 1518

Qy       1020 SerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGln 1039
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1519 TCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAG 1578

Qy       1040 ProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHis 1059
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1579 CCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCAC 1638

Qy       1060 CysPro 1061
              ||||||
Db       1639 TGTCCA 1644

RESULT 17
AAS72658
ID   AAS72658 standard; cDNA; 671 BP.
XX
AC   AAS72658;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #8462.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG08471.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 8462; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 671 BP; 207 A; 147 C; 148 G; 169 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.72e-87       Length:          671
Score:                  110.00         Matches:         110
Percent Similarity:     100.00%        Conservative:    0
Best Local Similarity:  100.00%        Mismatches:      0
Query Match:            7.70%          Indels:          0
DB:                     5              Gaps:            0

US-09-787-097-12 (1-1429) x AAS72658 (1-671)

Qy       1158 ThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyr 1177
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 ACTCTTCTTATTGACTATCAGTTCACTTTCAGTCTATCCCAGGAAGATGATCGCTATTAC 184

Qy       1178 ThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIle 1197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 ACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATC 244

Qy       1198 AsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThr 1217
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 AATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACC 304

Qy       1218 GlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSer 1237
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 CAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGT 364

Qy       1238 PheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrVal 1257
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 TTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTC 424

Qy       1258 SerAsnPheThrTrpProIleLysIleGln 1267
                |||||||||||||||||||||||||||||
Db        425 AGTAATTTCACCTGGCCCATCAAAATTCAG 454

RESULT 18
AAS02409
ID   AAS02409 standard; cDNA; 668 BP.
XX
AC   AAS02409;
XX
DT   18-JUL-2001  (first entry)
XX
DE   Human secreted protein, cDNA #15.
XX
KW   Human; secreted protein; immunogen; antibody; diagnosis;
KW   rheumatoid arthritis; hyperproliferative disorder; neoplasm;
KW   cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia;
KW   angiogenesis; Alzheimer's disease; bacterial infection; viral infection;
KW   fungal infection; corneal infection; wound healing; cell culture;
KW   epithelial cell proliferation; skin ageing; transplantation;
KW   tissue regeneration; chemotaxis; food additive; ss.
XX
OS   Homo sapiens.
XX
PN   WO200123546-A1.
XX
PD   05-APR-2001.
XX
PF   26-SEP-2000; 2000WO-US026323.
XX
PR   27-SEP-1999;   99US-0155805P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben S,  Komatsoulis GA;
XX
DR   WPI; 2001-266150/27.
DR   P-PSDB; AAU01736, AAU01774.
XX
PT   Nucleic acids encoding 37 human secreted polypeptides, useful for
PT   preventing, diagnosing and/or treating e.g. cancers, Parkinson's disease
PT   and diabetic retinopathy.
XX
PS   Disclosure; Page 420; 494pp; English.
XX
CC   The sequence encodes a human secreted protein of the invention. The
CC   polynucleotides, polypeptides and antibodies raised against them are used
CC   to prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC   rabbits, goats, horses, cats, dogs, chickens or sheep. The
CC   polynucleotides and antibodies are also used in diagnosing a pathological
CC   condition or susceptibility to a pathological condition. The antibodies
CC   can also be used in alleviating symptoms associated with the disorders
CC   and in diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
CC   immunosorbent assays (ELISA). Disorders which are diagnosed or treated
CC   include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
CC   disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
CC   e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
CC   angiogenesis, nervous system disorders e.g. Alzheimer's disease,
CC   infections caused by bacteria, viruses and fungi and ocular disorders
CC   e.g. corneal infection. The polypeptides can also be used to aid wound
CC   healing and epithelial cell proliferation, to prevent skin aging due to
CC   sunburn, to maintain organs before transplantation, for supporting cell
CC   culture of primary tissues, to regenerate tissues and in chemotaxis. The
CC   polypeptides can also be used as a food additive or preservative to
CC   increase or decrease storage capabilities
XX
SQ   Sequence 668 BP; 298 A; 113 C; 120 G; 137 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              3.48e-82        Length:         668
Score:                  104.00          Matches:        104
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            7.28%           Indels:         0
DB:                     4               Gaps:           0

US-09-787-097-12 (1-1429) x AAS02409 (1-668)

Qy       1164 GlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIleAsnPheVal 1183
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 CAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTG 70

Qy       1184 AlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPhe 1203
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTC 130

Qy       1204 AsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMet 1223
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 AACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATG 190

Qy       1224 ProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPhe 1243
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 CCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTT 250

Qy       1244 AspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpPro 1263
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 GATTTTCGCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCC 310

Qy       1264 IleLysIleGln 1267
                ||||||||||||
Db        311 ATCAAAATTCAG 322

RESULT 19
AAZ91917
ID   AAZ91917 standard; cDNA; 2419 BP.
XX
AC   AAZ91917;
XX
DT   08-JUN-2000  (first entry)
XX
DE   Murine mahogany protein coding sequence akm1003.
XX
KW   Mahogany gene; mouse; mg gene; regulatory defect; gene therapy; obesity;
KW   weight regulation; cell therapy; body weight disorder; cachexia;
KW   anorexia; hyperpigmentation; increased metabolic rate disorder;
KW   hyperphagia; Antiobesity; antianorexic; anticachexic; ss.
XX
OS   Mus sp.
XX
PN   WO200005373-A2.
XX
PD   03-FEB-2000.
XX
PF   21-JUL-1999;   99WO-US016484.
XX
PR   21-JUL-1998;   98US-0093630P.
PR   20-OCT-1998;   98US-0104978P.
PR   05-FEB-1999;   99US-00245041.
XX
PA   (MILL-) MILLENIUM PHARM INC.
XX
PI   Moore K,  Nagle DL;
XX
DR   WPI; 2000-195103/17.
DR   P-PSDB; AAY81804.
XX
PT   New human and murine mahogany genes, useful, e.g. for diagnosis and
PT   treatment of body weight disorders.
XX
```

PS Claim 1; Fig 8a; 188pp; English.
XX
CC This sequence represents a murine mahogany gene of the invention. The
CC mahogany genes are used: (i) to produce recombinant mahogany (mg)
CC proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming
CC therapeutics; (iii) as a source of diagnostic probes and primers for
CC detecting expression of mg genes or mutations, regulatory defects, in
CC this gene, or for isolation of related sequences; and (iv) in (cell-
CC based) gene therapy. (II) are used to raise specific antibodies (Ab); to
CC identify other (extra)cellular products involved in weight regulation,
CC and to screen for agents that disrupt interaction between (II) and other
CC macromolecules. The Ab are used to detect abnormal levels (or function)
CC of (II) (for diagnosis, prognosis or monitoring of treatment); to
CC evaluate (II)-expressing cells intended for cell therapy, and as
CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the
CC mg polypeptide) are used to identify agents (A) that modulate mg
CC activity. (A) are potentially useful for the treatment of body weight
CC disorders, particularly obesity, cachexia or anorexia, or other
CC conditions associated with the mg gene such as hyperpigmentation,
CC hyperphagia and disorders that result in increased metabolic rate
XX
SQ Sequence 2419 BP; 578 A; 567 C; 680 G; 594 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	8.11e-73	Length:	2419
Score:	94.00	Matches:	94
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	6.58%	Indels:	0
DB:	3	Gaps:	0

US-09-787-097-12 (1-1429) x AAZ91917 (1-2419)

```
Qy        513 AlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAsp 532
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1803 GCTCTGTACGTTCATGGTGGCTACAAGGCTTTCAGCGCCAACAAATACCGGCTTGCAGAT 1862

Qy        533 AspLeuTyrArgTyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPhe 552
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1863 GACCTCTACAGATACGATGTGGATACTCAGATGTGGACCATTCTTAAGGACAGCCGATTT 1922

Qy        553 PheArgTyrLeuHisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsn 572
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1923 TTCCGTTACTTGCATACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGAGGGAAC 1982

Qy        573 ThrHisAsnAspThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAla 592
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1983 ACACACAATGACACTTCCATGAGCCACGGTGCCAAATGCTTCTCCTCGGACTTCATGGCT 2042

Qy        593 TyrAspIleAlaCysAspArgTrpSerValLeuProArgPro 606
              ||||||||||||||||||||||||||||||||||||||
Db       2043 TATGACATTGCTTGTGACCGATGGTCAGTGCTTCCCAGACCT 2084
```

RESULT 20
AAL14686/c
ID AAL14686 standard; cDNA; 481 BP.
XX
AC AAL14686;
XX
DT 07-DEC-2001 (first entry)
XX
DE Human breast cancer expressed polynucleotide 7143.
XX
KW Human; breast cancer; cell marker; cytostatic; ss.
XX
OS Homo sapiens.
XX
PN WO200151628-A2.
XX
PD 19-JUL-2001.
XX
PF 10-JAN-2001; 2001WO-US000798.
XX
PR 14-JAN-2000; 2000US-0176077P.
PR 14-MAR-2000; 2000US-0189167P.
PR 24-MAR-2000; 2000US-0192099P.
PR 29-MAR-2000; 2000US-0193480P.
PR 15-MAY-2000; 2000US-0205230P.
PR 09-JUN-2000; 2000US-0211315P.
PR 25-JUL-2000; 2000US-0220534P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX
DR WPI; 2001-451856/48.
XX
PT New peptide useful as a marker for the diagnosis of breast cancer.
XX
PS Claim 1; Page 1292; 3695pp; English.
XX
CC The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterising treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity
XX
SQ Sequence 481 BP; 106 A; 133 C; 126 G; 116 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.07e-71	Length:	481
Score:	92.00	Matches:	92
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	6.44%	Indels:	0
DB:	4	Gaps:	0

US-09-787-097-12 (1-1429) x AAL14686 (1-481)

```
Qy        901 ThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArg 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 ACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGG 218

Qy        921 ProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGly 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 CCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGA 158

Qy        941 AspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAsp 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 GATTGCACCAGCGGCAGCTCTGAGTGTATGTGGTGCAGCAACATGAAGCAGTGTGTGGAC 98

Qy        961 SerAsnAlaTyrValAlaSerProProPheGlyGlnCysMetGluTrpTyrThrMetSer 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         97 TCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGC 38

Qy        981 ThrCysProProGluAsnCysSerGlyTyrCysThr 992
              ||||||||||||||||||||||||||||||||||
Db         37 ACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACC 2
```

RESULT 21
AAX84568
ID AAX84568 standard; cDNA to mRNA; 1012 BP.
XX
AC AAX84568;
XX
DT 13-SEP-1999 (first entry)
XX
DE Biorhythm marker gene.
XX
KW Biorhythm marker gene; circadian rhythm; ss.
XX
OS Rattus sp.

```
XX
PN   JP11169185-A.
XX
PD   29-JUN-1999.
XX
PF   12-DEC-1997;   97JP-00362890.
XX
PR   12-DEC-1997;   97JP-00362890.
XX
PA   (AGEN ) AGENCY OF IND SCI & TECHNOLOGY.
PA   (SAKA ) OTSUKA PHARM CO LTD.
XX
DR   WPI; 1999-422623/36.
DR   P-PSDB; AAY22205.
XX
PT   Structure and function of a new biorrhythm marker gene - and protein
PT   encoded by it.
XX
PS   Claim 3; Page 12; 23pp; Japanese.
XX
CC   This sequence encodes the biorhythm marker protein of the invention. The
CC   protein is also referred to as a circadian rhythm protein
XX
SQ   Sequence 1012 BP; 263 A; 250 C; 257 G; 242 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              2.13e-71        Length:         1012
Score:                  92.00           Matches:        92
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            6.44%           Indels:         0
DB:                     2               Gaps:           0

US-09-787-097-12 (1-1429) x AAX84568 (1-1012)

Qy        849 SerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 TCCAAGCTCACTCTGACTCCATGGGTTGGTCTTCGGAAGATCAATGTATCTTACTGGTGC 372

Qy        869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        373 TGGGAGGATATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGT 432

Qy        889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        433 GATGCCGGCTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGACTGAAGGCTGCAACC 492

Qy        909 CysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSerAlaLysGln 928
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        493 TGCATCAACCCACTCAATGGAAGTGTTTGTGAAAGGCCTGCAAACCACAGTGCCAAGCAG 552

Qy        929 CysArgThrProCysAlaLeuArgThrAlaCysGly 940
              ||||||||||||||||||||||||||||||||||||
Db        553 TGCCGGACACCCTGTGCCCTGCGGACAGCATGTGGT 588

RESULT 22
ABK38928/c
ID   ABK38928 standard; cDNA; 246 BP.
XX
AC   ABK38928;
XX
DT   21-MAY-2002  (first entry)
XX
DE   cDNA encoding lung tumour protein clone R0127:F03.
XX
KW   Lung tumour; cancer; T cell; immune response stimulator; cytostatic;
KW   gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200204514-A2.
XX
PD   17-JAN-2002.
XX
PF   10-JUL-2001; 2001WO-US022058.
XX
PR   11-JUL-2000; 2000US-00614124.
PR   29-AUG-2000; 2000US-00651563.
PR   08-SEP-2000; 2000US-00658824.
PR   26-SEP-2000; 2000US-00671325.
PR   06-OCT-2000; 2000US-00677419.
PR   30-OCT-2000; 2000US-00702705.
PR   13-DEC-2000; 2000US-00736457.
PR   03-MAY-2001; 2001US-00849626.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Wang T,  Watanabe Y,  Henderson RA,  Johnson JC,  Retter MW;
PI   Marnerakis M,  Carter D,  Fanger GR,  Vedvick TS,  Bangur CS,  Mcnabb A;
PI   Wang A,  Fanger N,  Switzer A,  Mcneill PD,  Clapper JD;
XX
DR   WPI; 2002-164634/21.
XX
PT   Novel polynucleotide encoding a lung tumor polypeptide useful for
PT   stimulating and/or expanding T cells specific for a tumor protein.
XX
PS   Claim 1; SEQ ID NO 966; 223pp; English.
XX
CC   The invention describes an isolated polynucleotide and polypeptide useful
CC   for stimulating and/or expanding T cells specific for a tumour protein
CC   for determining the presence of a cancer in a patient. A composition
CC   containing the polynucleotide and/or polypeptide is useful for treating a
CC   lung cancer in a patient. The polypeptide is useful for removing tumour
CC   cells from a biological sample. The polynucleotide is also useful as
CC   probe or primer to detect the level of mRNA encoding a tumour protein.
CC   This sequence encodes a lung tumour associated protein or protein
CC   fragment, described in the method of the invention. Note: The sequence
CC   data for this patent did not form part of the printed specification, but
CC   was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 246 BP; 63 A; 61 C; 60 G; 62 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              4.04e-63        Length:         246
Score:                  82.00           Matches:        82
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            5.74%           Indels:         0
DB:                     6               Gaps:           0

US-09-787-097-12 (1-1429) x ABK38928 (1-246)

Qy        839 LeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGly 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy        859 LeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSer 878
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy        879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGlu 898
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67

Qy        899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy        919 GluArg 920
              ||||||
Db          6 GAAAGG 1

RESULT 23
ACA11257/c
ID   ACA11257 standard; cDNA; 246 BP.
```

XX
AC ACA11257;
XX
DT 05-JUN-2003 (first entry)
XX
DE Human lung adenocarcinoma library cDNA SEQ ID 966.
XX
KW Human; lung cancer; ss; lung tumour; cytostatic; vaccine;
KW T cell expansion; CD4; CD8.
XX
OS Homo sapiens.
XX
PN US2002197669-A1.
XX
PD 26-DEC-2002.
XX
PF 03-MAY-2001; 2001US-00849626.
XX
PR 13-DEC-2000; 2000US-00736457.
XX
PA (BANG/) BANGUR C S.
PA (FANG/) FANGER G R.
PA (WANG/) WANG A.
PA (WANG/) WANG T.
PA (SWIT/) SWITZER A P.
PA (MCNE/) MCNEILL P D.
PA (CLAP/) CLAPPER J D.
XX
PI Bangur CS, Fanger GR, Wang A, Wang T, Switzer AP, Mcneill PD;
PI Clapper JD;
XX
DR WPI; 2003-352750/33.
XX
PT Novel lung cancer polynucleotide encoding lung cancer protein, useful for
PT detecting the presence of lung cancer in a patient, and in pharmaceutical
PT compositions, e.g. vaccines, for treating lung cancer.
XX
PS Example 1; Page; 72pp; English.
XX
CC The invention relates to a polynucleotide encoding a lung tumour protein,
CC comprising a sequence selected from any of the 14 sequences mentioned in
CC the specification, or a sequence (S2) mentioned in specification,
CC complement of S1, sequences consisting of at least 20 contiguous residues
CC of S1, sequences that hybridise to S1, sequences having 75%, preferably
CC 90%, identity to S1, or degenerate variants of S1. Also included are an
CC isolated polypeptide (comprising a sequence (S3) selected from any one of
CC the 4 amino acid sequences mentioned in the specification, a sequence
CC encoded by the polynucleotide, or sequences having at least 70%,
CC preferably 90%, identity to a sequence encoded by the polynucleotide), an
CC expression vector comprising the polynucleotide operably linked to an
CC expression control sequence, a host cell transformed or transfected with
CC the vector, an isolated antibody (or its antigen-binding fragment) that
CC specifically binds to the polypeptide, detecting the presence of a cancer
CC in a patient, a fusion protein comprising the polypeptide, an
CC oligonucleotide that hybridises to S1 under moderately stringent
CC conditions, stimulating and/or expanding T cells specific for a tumour
CC protein (comprising contacting T cells with the polynucleotide, protein
CC or antigen-presenting cells, under conditions and for a time sufficient
CC to permit the stimulation and/or expansion of T cells) and inhibiting the
CC development of a cancer in a patient (by incubating CD4^+ and/or CD8^+ T
CC cells isolated from a patient with the polynucleotide, protein or antigen
CC presenting cells that express the polynucleotide, such that T cells
CC proliferate, administering to the patient an effective amount of the
CC proliferated T cells, and thus inhibiting the development of a cancer in
CC the patient. The polynucleotide, protein ands cells are useful in a
CC composition for stimulating an immune response in a patient, and for
CC treating a cancer in a patient (particularly lung cancer). The
CC oligonucleotide is useful for determining the presence of a cancer in a
CC patient. The protein and oligonucleotides are useful in pharmaceutical
CC compositions, e.g. vaccines. The polynucleotide is also useful as a probe
CC or primer for nucleic acid hybridisation, and in the design and
CC preparation of ribozyme molecules for inhibiting expression of tumour
CC polypeptides and proteins in tumour cells. An amplified portion of the
CC polynucleotide is useful for isolating a full-length gene from a suitable
CC library. The present sequence is a cDNA (full length, extended or
CC partial) isolated from a library derived from lung tumour/cancer cells.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO at seqdata.uspto.gov/sequence.html?DocId=20020197669
XX
SQ Sequence 246 BP; 63 A; 61 C; 60 G; 62 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4.04e-63	Length:	246
Score:	82.00	Matches:	82
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.74%	Indels:	0
DB:	7	Gaps:	0

US-09-787-097-12 (1-1429) x ACA11257 (1-246)

```
Qy        839 LeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGly 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy        859 LeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSer 878
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy        879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGlu 898
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67

Qy        899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy        919 GluArg 920
              ||||||
Db          6 GAAAGG 1
```

RESULT 24
ACA02443/c
ID ACA02443 standard; cDNA; 246 BP.
XX
AC ACA02443;
XX
DT 22-MAY-2003 (first entry)
XX
DE Lung cancer therapyand diagnosis associated cDNA #932.
XX
KW Lung cancer; cytostatic; vaccine; gene therapy; cancer; gene; ss.
XX
OS Homo sapiens.
XX
PN US2002172952-A1.
XX
PD 21-NOV-2002.
XX
PF 10-JUL-2001; 2001US-00902941.
XX
PR 30-JUN-1999; 99US-00346492.
PR 15-OCT-1999; 99US-00419356.
PR 17-DEC-1999; 99US-00466867.
PR 30-DEC-1999; 99US-00476300.
PR 06-MAR-2000; 2000US-00519642.
PR 22-MAR-2000; 2000US-00533077.
PR 10-APR-2000; 2000US-00546259.
PR 27-APR-2000; 2000US-00560406.
PR 05-JUN-2000; 2000US-00589184.
PR 11-JUL-2000; 2000US-00614124.
PR 29-AUG-2000; 2000US-00651563.
PR 08-SEP-2000; 2000US-00658824.
PR 26-SEP-2000; 2000US-00671325.
PR 06-OCT-2000; 2000US-00677419.

```
PR   30-OCT-2000; 2000US-00702705.
PR   13-DEC-2000; 2000US-00736457.
PR   03-MAY-2001; 2001US-00849626.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Henderson RA,  Wang T,  Watanabe Y,  Johnson JC,  Retter MW;
PI   Durham M,  Carter D,  Fanger GR,  Vedvick TS,  Bangur CS,  Mcnabb A;
XX
DR   WPI; 2003-328427/31.
XX
PT   New polynucleotide, useful for preparing a composition for treating or
PT   inhibiting development of cancer, e.g. lung cancer.
XX
PS   Example 1; SEQ ID NO 966; 82pp; English.
XX
CC   The invention describes an isolated polynucleotide comprising one of 32
CC   sequences, complement or degenerate variants of them. The polynucleotide
CC   is useful for preparing a composition e.g. a vaccine or for gene therapy,
CC   for treating or inhibiting development of cancer, e.g. lung cancer. This
CC   sequence represents a polynucleotide associated with the compositions and
CC   methods for the therapy and diagnosis of lung cancer
XX
SQ   Sequence 246 BP; 63 A; 61 C; 60 G; 62 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                4.04e-63        Length:          246
Score:                    82.00           Matches:         82
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              5.74%           Indels:          0
DB:                       7               Gaps:            0

US-09-787-097-12 (1-1429) x ACA02443 (1-246)

Qy        839 LeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGly 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy        859 LeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSer 878
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy        879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGlu 898
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67

Qy        899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy        919 GluArg 920
              ||||||
Db          6 GAAAGG 1

RESULT 25
AAL23548/c
ID   AAL23548 standard; cDNA; 482 BP.
XX
AC   AAL23548;
XX
DT   07-DEC-2001  (first entry)
XX
DE   Human breast cancer expressed polynucleotide 16005.
XX
KW   Human; breast cancer; cell marker; cytostatic; ss.
XX
OS   Homo sapiens.
XX
PN   WO200151628-A2.
XX
PD   19-JUL-2001.
XX
PF   10-JAN-2001; 2001WO-US000798.
XX
PR   14-JAN-2000; 2000US-0176077P.
PR   14-MAR-2000; 2000US-0189167P.
PR   24-MAR-2000; 2000US-0192099P.
PR   29-MAR-2000; 2000US-0193480P.
PR   15-MAY-2000; 2000US-0205230P.
PR   09-JUN-2000; 2000US-0211315P.
PR   25-JUL-2000; 2000US-0220534P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Lillie J,  Xu Y,  Wang Y,  Steinmann K;
XX
DR   WPI; 2001-451856/48.
XX
PT   New peptide useful as a marker for the diagnosis of breast cancer.
XX
PS   Claim 1; Page 2922; 3695pp; English.
XX
CC   The invention relates to human breast cancer expressed polynucleotides
CC   (AAL07544-AAL26789) and methods of assessing whether a patient is
CC   afflicted with breast cancer by examining the correlation between the
CC   expression of certain markers and the cancerous state of breast cells.
CC   The polynucleotides and encoded polypeptides are potential markers for
CC   detecting, diagnosing, monitoring, characterising treating and
CC   potentially preventing breast cancer. The polynucleotides and encoded
CC   polypeptides are also useful for isolating compounds with cytostatic
CC   activity
XX
SQ   Sequence 482 BP; 106 A; 133 C; 126 G; 117 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                3.41e-60        Length:          482
Score:                    79.00           Matches:         92
Percent Similarity:       98.92%          Conservative:    0
Best Local Similarity:    98.92%          Mismatches:      0
Query Match:              5.53%           Indels:          1
DB:                       4               Gaps:            0

US-09-787-097-12 (1-1429) x AAL23548 (1-482)

Qy        901 ThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArg 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 ACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGG 219

Qy        921 ProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAla-CysGl 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        218 CCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCAATGTGG 159

Qy        940 yAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAs 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 AGATTGCACCAGCGGCAGCTCTGAGTGTATGTGGTGCAGCAACATGAAGCAGTGTGTGGA 99

Qy        960 pSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSe 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         98 CTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAG 39

Qy        980 rThrCysProProGluAsnCysSerGlyTyrCysThr 992
              ||||||||||||||||||||||||||||||||||||||
Db         38 CACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACC 2

RESULT 26
ABZ72040
ID   ABZ72040 standard; DNA; 207433 BP.
XX
AC   ABZ72040;
XX
DT   03-APR-2003  (first entry)
XX
DE   Gene 216 H194BAC1098L22 nucleotide sequence SEQ ID NO 5.
XX
KW   Human; Gene 216; chromosome 20p13-p12; antiasthmatic; anorectic;
```

KW antiinflammatory; gastrointestinal; gene therapy; vaccine; asthma;
KW obesity; inflammatory bowel disease; promoter; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200178894-A2.
XX
PD 25-OCT-2001.
XX
PF 13-APR-2001; 2001WO-US012245.
XX
PR 13-APR-2000; 2000US-00548797.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Keith T;
XX
DR WPI; 2001-639428/73.
DR P-PSDB; ABR00926.
XX
PT Isolated genes (Gene 216) from human chromosome 20p13-p12 and the
PT proteins they encode, useful for the prevention, diagnosis and treatment
PT of asthma, obesity and inflammatory bowel disease.
XX
PS Example 4; Fig 7; 520pp; English.
XX
CC The invention relates to isolated genes (Gene 216) from human chromosome
CC 20p13-p12 and the proteins they encode. The nucleic acids and proteins
CC may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate Gene 216 expression. For example, the
CC nucleic acids (or vectors) and proteins may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of gene 216 by expressing
CC inactive proteins or to supplement the patients own production of Gene
CC 216 proteins. Additionally, the nucleic acids may be used to produce the
CC secreted Gene 216 protein, by inserting the nucleic acids into a host
CC cell and culturing the cell to express the protein. The nucleic acids and
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar nucleic acid
CC sequences in samples and therefore which patients may be in need of
CC restorative therapy. The Gene 216 protein may also be used as antigens in
CC the production of antibodies against Gene 216 and in assays to identify
CC modulators of Gene 216 expression and activity. The anti-Gene 216
CC antibodies and antagonists may also be used to down regulate expression
CC and activity. The anti-Gene 216 antibodies may also be used as diagnostic
CC agents for detecting the presence of Gene 216 proteins in samples (e.g.
CC by enzyme linked immunosorbant assay or ELISA). Disorders that may be
CC prevented, diagnosed and/or treated by the above methods include, for
CC example asthma, obesity and inflammatory bowel disease. The present
CC sequence is that of the Gene 216 genomic nucleic acid sequence, promoter
CC or enhancer
XX
SQ Sequence 207433 BP; 52775 A; 51290 C; 51698 G; 51670 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	5.56e-56	Length:	207433
Score:	77.00	Matches:	77
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.39%	Indels:	0
DB:	5	Gaps:	0

US-09-787-097-12 (1-1429) x ABZ72040 (1-207433)

```
Qy       985 GluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrp 1004
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9815 GAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGG 9874

Qy      1005 CysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyPro 1024
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9875 TGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCA 9934

Qy      1025 ValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSer 1044
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9935 GTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCC 9994

Qy      1045 SerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysPro 1061
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9995 AGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCA 10045
```

RESULT 27
ABX74891
ID ABX74891 standard; DNA; 207433 BP.
XX
AC ABX74891;
XX
DT 07-APR-2003 (first entry)
XX
DE BAC1098L22 DNA sequence.
XX
KW Gene 216; antiasthmatic; antiinflammatory; ss; anorectic;
KW chromosome 20p13-p12; single nucleotide polymorphism; SNP; gene therapy;
KW respiratory disease; asthma; obesity; bronchial hyper-responsiveness;
KW chronic obstructive pulmonary disease;
KW adult respiratory distress syndrome; inflammatory bowel syndrome.
XX
OS Synthetic.
XX
PN WO200283077-A2.
XX
PD 24-OCT-2002.
XX
PF 15-APR-2002; 2002WO-US012063.
XX
PR 13-APR-2001; 2001US-00834597.
PR 13-APR-2001; 2001WO-US012245.
XX
PA (SCHE) SCHERING CORP.
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Keith T, Little RD, Van Eerdewegh P, Dupuis J, Del Mastro RG;
PI Simon J, Allen K, Pandit S;
XX
DR WPI; 2003-092960/08.
XX
PT New isolated gene 216 nucleic acids, useful for diagnosing, preventing or
PT treating a disorder, such as asthma, bronchial hyper-responsiveness,
PT chronic obstructive pulmonary disease, obesity or inflammatory bowel
PT syndrome.
XX
PS Example 6; Fig 7; 650pp; English.
XX
CC This invention relates to a novel isolated nucleic acid, gene 216,
CC identified from human chromosome 20p13-p12. The invention also discloses
CC regions of the 216 gene that contain single nucleotide polymorphisms
CC (SNP's) which may be used as markers for disease susceptibility or
CC severity. The nucleotides of the invention may have antiasthmatic,
CC antiinflammatory or anorectic activities and may be used in gene therapy.
CC The nucleic acids, antibodies or its fragments are useful for diagnosing,
CC preventing or treating a disorder, such as respiratory diseases (e.g.
CC asthma, bronchial hyper-responsiveness, chronic obstructive pulmonary
CC disease or adult respiratory distress syndrome), obesity, or inflammatory
CC bowel syndrome. The nucleic acids are also useful for identifying
CC increased susceptibility of a subject to the disorders mentioned. The
CC nucleic acids can also be used as primers and templates for the
CC recombinant production of disorder-associated peptides or polypeptides,
CC for chromosome and gene mapping, or for tissue distribution studies. The
CC present sequence represents a gene 216 cDNA sequence used in the scope of
CC the invention
XX
SQ Sequence 207433 BP; 52775 A; 51289 C; 51698 G; 51671 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	5.56e-56	Length:	207433
Score:	77.00	Matches:	77
Percent Similarity:	100.00%	Conservative:	0

```
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            5.39%          Indels:        0
DB:                     7              Gaps:          0

US-09-787-097-12 (1-1429) x ABX74891 (1-207433)

Qy        985 GluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrp 1004
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9815 GAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGG 9874

Qy       1005 CysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyPro 1024
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9875 TGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCA 9934

Qy       1025 ValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSer 1044
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9935 GTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCC 9994

Qy       1045 SerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysPro 1061
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       9995 AGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCA 10045

RESULT 28
AAZ91918
ID   AAZ91918 standard; cDNA; 1051 BP.
XX
AC   AAZ91918;
XX
DT   08-JUN-2000  (first entry)
XX
DE   Murine mahogany protein coding sequence akm1004.
XX
KW   Mahogany gene; mouse; mg gene; regulatory defect; gene therapy; obesity;
KW   weight regulation; cell therapy; body weight disorder; cachexia;
KW   anorexia; hyperpigmentation; increased metabolic rate disorder;
KW   hyperphagia; Antiobesity; antianorexic; anticachexic; ss.
XX
OS   Mus sp.
XX
PN   WO200005373-A2.
XX
PD   03-FEB-2000.
XX
PF   21-JUL-1999;   99WO-US016484.
XX
PR   21-JUL-1998;   98US-0093630P.
PR   20-OCT-1998;   98US-0104978P.
PR   05-FEB-1999;   99US-00245041.
XX
PA   (MILL-) MILLENIUM PHARM INC.
XX
PI   Moore K,  Nagle DL;
XX
DR   WPI; 2000-195103/17.
DR   P-PSDB; AAY81805.
XX
PT   New human and murine mahogany genes, useful, e.g. for diagnosis and
PT   treatment of body weight disorders.
XX
PS   Claim 1; Fig 9a; 188pp; English.
XX
CC   This sequence represents a murine mahogany gene of the invention. The
CC   mahogany genes are used: (i) to produce recombinant mahogany (mg)
CC   proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming
CC   therapeutics; (iii) as a source of diagnostic probes and primers for
CC   detecting expression of mg genes or mutations, regulatory defects, in
CC   this gene, or for isolation of related sequences; and (iv) in (cell-
CC   based) gene therapy. (II) are used to raise specific antibodies (Ab); to
CC   identify other (extra)cellular products involved in weight regulation,
CC   and to screen for agents that disrupt interaction between (II) and other
CC   macromolecules. The Ab are used to detect abnormal levels (or function)
CC   of (II) (for diagnosis, prognosis or monitoring of treatment); to
CC   evaluate (II)-expressing cells intended for cell therapy, and as
CC   therapeutic mg inhibitors. Cells that express the mg gene (or contain the
CC   mg polypeptide) are used to identify agents (A) that modulate mg
CC   activity. (A) are potentially useful for the treatment of body weight
CC   disorders, particularly obesity, cachexia or anorexia, or other
CC   conditions associated with the mg gene such as hyperpigmentation,
CC   hyperphagia and disorders that result in increased metabolic rate
XX
SQ   Sequence 1051 BP; 306 A; 243 C; 313 G; 189 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              3.17e-57       Length:        1051
Score:                  76.00          Matches:       111
Percent Similarity:     98.23%         Conservative:  0
Best Local Similarity:  98.23%         Mismatches:    1
Query Match:            5.32%          Indels:        2
DB:                     3              Gaps:          0

US-09-787-097-12 (1-1429) x AAZ91918 (1-1051)

Qy         87 AlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGluCysAspArgProCysVal 106
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        463 GCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAAGGAATGTGACCGGCCGTGTGTC 522

Qy        107 AsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCys-ProAlaGlyTrpValGl 126
              |||||||||||||||||||||||||||||||||||||||||| ||| ||||||||||||
Db        523 AACGGCGGCCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCACG-GGCTGGGTGGG 581

Qy        126 yGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPheValTh 146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        582 CGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGCTCTTCTGGATTTGTAAC 641

Qy        146 rAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeuIleGluGlyGlnPr 166
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        642 AGATGGACCTGGGAATTATAAATATAAGACGAAGTGCACATGGCTCATTGAAGGACAGCC 701

Qy        166 oAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLe 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        702 AAATAGAATAATGAGACTTCGCTTCAACCATTTTGCTACAGAATGTAGCTGGGACCATTT 761

Qy        186 uTyrValTyrAspGlyAspSerIleTyrAlaProLeu 198
              |||||||||||||||||||||||||||||||||||
Db        762 ATATGTTTATGATGGGGACTCAATCTACGCACCTCTG 798

RESULT 29
AAZ91925
ID   AAZ91925 standard; DNA; 90050 BP.
XX
AC   AAZ91925;
XX
DT   08-JUN-2000  (first entry)
XX
DE   Wild type (C57BL/6J) murine mahogany protein genomic sequence c110/111.
XX
KW   Mahogany gene; mouse; mg gene; regulatory defect; gene therapy; obesity;
KW   weight regulation; cell therapy; body weight disorder; cachexia;
KW   anorexia; hyperpigmentation; increased metabolic rate disorder;
KW   hyperphagia; Antiobesity; antianorexic; anticachexic; ds.
XX
OS   Mus sp.
XX
PN   WO200005373-A2.
XX
PD   03-FEB-2000.
XX
PF   21-JUL-1999;   99WO-US016484.
XX
PR   21-JUL-1998;   98US-0093630P.
PR   20-OCT-1998;   98US-0104978P.
PR   05-FEB-1999;   99US-00245041.
XX
PA   (MILL-) MILLENIUM PHARM INC.
XX
PI   Moore K,  Nagle DL;
```

XX
DR WPI; 2000-195103/17.
XX
PT New human and murine mahogany genes, useful, e.g. for diagnosis and
PT treatment of body weight disorders.
XX
PS Example; Fig 3d; 188pp; English.
XX
CC This sequence represents a murine mahogany gene of the invention. The
CC mahogany genes are used: (i) to produce recombinant mahogany (mg)
CC proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming
CC therapeutics; (iii) as a source of diagnostic probes and primers for
CC detecting expression of mg genes or mutations, regulatory defects, in
CC this gene, or for isolation of related sequences; and (iv) in (cell-
CC based) gene therapy. (II) are used to raise specific antibodies (Ab); to
CC identify other (extra)cellular products involved in weight regulation,
CC and to screen for agents that disrupt interaction between (II) and other
CC macromolecules. The Ab are used to detect abnormal levels (or function)
CC of (II) (for diagnosis, prognosis or monitoring of treatment); to
CC evaluate (II)-expressing cells intended for cell therapy, and as
CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the
CC mg polypeptide) are used to identify agents (A) that modulate mg
CC activity. (A) are potentially useful for the treatment of body weight
CC disorders, particularly obesity, cachexia or anorexia, or other
CC conditions associated with the mg gene such as hyperpigmentation,
CC hyperphagia and disorders that result in increased metabolic rate
XX
SQ Sequence 90050 BP; 25307 A; 18193 C; 19454 G; 27096 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	8.88e-53	Length:	90050
Score:	73.00	Matches:	73
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.11%	Indels:	0
DB:	3	Gaps:	0

US-09-787-097-12 (1-1429) x AAZ91925 (1-90050)

```
Qy        849 SerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      28996 TCCAAGCTCACTCTGACTCCATGGGTTGGTCTTCGGAAGATCAATGTGTCTTACTGGTGC 29055

Qy        869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      29056 TGGGAGGATATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGT 29115

Qy        889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      29116 GATGCTGGCTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGATTAAAGGCTGCAACC 29175

Qy        909 CysIleAsnProLeuAsnGlySerValCysGluArgPro 921
              |||||||||||||||||||||||||||||||||||||||
Db      29176 TGCATCAACCCTCTCAATGGCAGCGTCTGTGAAAGGCCT 29214
```

RESULT 30
AAI43425
ID AAI43425 standard; DNA; 531 BP.
XX
AC AAI43425;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #12111 used to measure gene expression in human placenta sample.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
PS Claim 25; SEQ ID NO 12111; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders
XX
SQ Sequence 531 BP; 145 A; 82 C; 107 G; 197 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.63e-51	Length:	531
Score:	69.00	Matches:	69
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	4.83%	Indels:	0
DB:	4	Gaps:	0

US-09-787-097-12 (1-1429) x AAI43425 (1-531)

```
Qy        414 ThrGlyAsnValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeu 433
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 ACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTG 245

Qy        434 LeuThrProLysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThr 453
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 TTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACA 305

Qy        454 LeuLysAsnGlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyr 473
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 CTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATAT 365

Qy        474 IleSerAsnValGlnGluTyrAspLeu 482
              |||||||||||||||||||||||||||
Db        366 ATAAGCAATGTGCAGGAATATGATTTG 392
```

RESULT 31
ABS11523
ID ABS11523 standard; DNA; 531 BP.
XX
AC ABS11523;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human genome-derived single exon probe from lung SEQ ID No 11514.
XX
KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;

KW primary ciliary dyskinesis; pulmonary hypertension;
KW hyaline membrane disease.
XX
OS Homo sapiens.
XX
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US000665.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2002-114183/15.
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.
XX
PS Claim 1; SEQ ID NO 11514; 634pp; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC ; the novel set of probes which hybridise at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a single exon probe of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 531 BP; 145 A; 82 C; 107 G; 197 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.63e-51 Length: 531

Score:	69.00	Matches:	69
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	4.83%	Indels:	0
DB:	6	Gaps:	0

US-09-787-097-12 (1-1429) x ABS11523 (1-531)

```
Qy      414 ThrGlyAsnValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeu 433
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      186 ACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTG 245

Qy      434 LeuThrProLysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThr 453
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      246 TTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACA 305

Qy      454 LeuLysAsnGlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyr 473
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      306 CTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATAT 365

Qy      474 IleSerAsnValGlnGluTyrAspLeu 482
            |||||||||||||||||||||||||||
Db      366 ATAAGCAATGTGCAGGAATATGATTTG 392
```

RESULT 32
AAI56405
ID AAI56405 standard; DNA; 193 BP.
XX
AC AAI56405;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #25091 used to measure gene expression in human placenta sample.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
PS Claim 25; SEQ ID NO 25091; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders
XX
SQ Sequence 193 BP; 50 A; 36 C; 52 G; 55 T; 0 U; 0 Other;

Alignment Scores:

```
Pred. No.:              2.74e-47        Length:         193
Score:                  64.00           Matches:        64
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            4.48%           Indels:         0
DB:                     4               Gaps:           0

US-09-787-097-12 (1-1429) x AAI56405 (1-193)

Qy        416 AsnValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThr 435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 AATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACC 61

Qy        436 ProLysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLys 455
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAG 121

Qy        456 AsnGlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSer 475
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 AATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGC 181

Qy        476 AsnValGlnGlu 479
              ||||||||||||
Db        182 AATGTGCAGGAA 193

RESULT 33
ABS23881
ID   ABS23881 standard; DNA; 193 BP.
XX
AC   ABS23881;
XX
DT   19-AUG-2002  (first entry)
XX
DE   Human genome-derived single exon probe ORF from lung SEQ ID No 23872.
XX
KW   Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease; open reading frame; ORF.
XX
OS   Homo sapiens.
XX
PN   WO200186003-A2.
XX
PD   15-NOV-2001.
XX
PF   30-JAN-2001; 2001WO-US000665.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2002-114183/15.
XX
PT   Spatially-addressable set of single exon nucleic acid probes, used to
PT   measure gene expression in human lung samples.
XX
PS   Claim 4; SEQ ID NO 23872; 634pp; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human lung comprising single exon nucleic acid probes having one of
CC   12614 nucleic acid sequences mentioned in the specification, or their
CC   complements or the 12387 open reading frames derived from the 12614
CC   probes. Also included are a microarray comprising the novel set of probes
CC   ; the novel set of probes which hybridise at high stringency to a nucleic
CC   acid expressed in the human lung; measuring gene expression in a sample
CC   derived from human lung, comprising (a) contacting the array with a
CC   collection of detectably labeled nucleic acids derived from human lung
CC   mRNA, and (b) measuring the label detectably bound to each probe of the
CC   array; identifying exons in a eukaryotic genome, comprising (a)
CC   algorithmically predicting at least one exon from genomic sequences of
CC   the eukaryote; and (b) detecting specific hybridisation of detectably
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC   having a fragment identical to the predicted exon, the probe is included
CC   in the above mentioned microarray; assigning exons to a single gene,
CC   comprising (a) identifying exons from genomic sequence by the method
CC   above and (b) measuring the expression of each of the exons in several
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern of
CC   expression of the exons in the tissues and/or cell types indicates that
CC   the exons should be assigned to a single gene; a peptide comprising one
CC   of 12011 sequences, mentioned in the specification, or encoded by the
CC   probes/open reading frames (ORF). The probes are used for gene expression
CC   analysis, and for identifying exons in a gene, particularly using human
CC   lung derived mRNA and for the study of lung diseases such as asthma, lung
CC   cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC   disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC   tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC   Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC   histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC   Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC   dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC   present sequence is a single exon probe open reading frame of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 193 BP; 50 A; 36 C; 52 G; 55 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              2.74e-47        Length:         193
Score:                  64.00           Matches:        64
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            4.48%           Indels:         0
DB:                     6               Gaps:           0

US-09-787-097-12 (1-1429) x ABS23881 (1-193)

Qy        416 AsnValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThr 435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 AATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACC 61

Qy        436 ProLysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLys 455
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAG 121

Qy        456 AsnGlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSer 475
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 AATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGC 181

Qy        476 AsnValGlnGlu 479
              ||||||||||||
Db        182 AATGTGCAGGAA 193

RESULT 34
AAA45505
ID   AAA45505 standard; cDNA; 292 BP.
XX
AC   AAA45505;
XX
DT   21-AUG-2000  (first entry)
```

XX
DE Human secreted expressed sequence tag SEQ ID NO:2080.
XX
KW Human; mouse; chicken; rat; secreted expressed sequence tag; sEST;
KW expressed sequence tag; EST; probe; chemotactic; proliferative;
KW immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;
KW thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;
KW antiviral; antidiabetic; antiasthmatic; vulnerary; antiparkinsonian;
KW antiulcer; osteopathic; neuroprotective; nootropic; antipsoriatic;
KW cerebroprotective; anticonvulsant; antidepressant; gene therapy; vaccine;
KW autoimmune disorder; multiple sclerosis; allergic condition;
KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;
KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;
KW central nervous system disorder; Alzheimer's disease; stroke;
KW Parkinson's disease; Huntington's disease; coagulation disorder;
KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease; tumour;
KW infection; depression; psoriasis; ss.
XX
OS Homo sapiens.
XX
PN WO200021991-A1.
XX
PD 20-APR-2000.
XX
PF 15-OCT-1999; 99WO-US024206.
XX
PR 15-OCT-1998; 98US-0104436P.
XX
PA (GEMY) GENETICS INST INC.
XX
PI Jacobs K, Mccoy JM, Lavallie ER, Collins-Racie LA, Evans C;
PI Merberg D, Treacy M, Bowman MR;
XX
DR WPI; 2000-317938/27.
XX
PT Isolated polynucleotides, and encoded proteins, comprising secreted
PT expressed sequence tags (sESTs), useful for treating various disorders
PT such as autoimmune, infectious, and central nervous system disorders.
XX
PS Claim 1; Page 724; 803pp; English.
XX
CC AAA43426 to AAA45925 represent specifically claimed secreted expressed
CC sequence tags (sESTs), isolated from human, mouse, chicken and rat tissue
CC sources. The sESTs can have a range of activities depending on the
CC tissues they were isolated from. The activities include: chemotactic;
CC proliferative; immunomodulatory; haematopoietic; chemokinetic; analgesic;
CC haemostatic; thrombolytic; antiinflammatory; cytostatic; antibacterial;
CC antifungal; antiviral; antidiabetic; antiasthmatic; vulnerary; antiulcer;
CC osteopathic; neuroprotective; nootropic; antiparkinsonian; antipsoriatic;
CC cerebroprotective; anticonvulsant; and antidepressant. The sESTs can be
CC used for gene therapy and in vaccines. The sESTs are useful as probes for
CC the identification and isolation of full-length cDNAs and genomic DNA
CC molecules which correspond to the sESTs. Proteins encoded by the sESTs
CC are useful in assays for determining biological activity and raising
CC antibodies. They may be useful for treatment of autoimmune disorders
CC (multiple sclerosis, insulin dependent diabetes), allergic conditions
CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,
CC osteoporosis, osteoarthritis, central nervous system disorders
CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation
CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's
CC disease), tumours, bacterial, fungal or viral infections, depression and
CC psoriasis. AAA45926 to AAA45931 represent linker variants which are given
CC in the exemplification of the present invention
XX
SQ Sequence 292 BP; 77 A; 58 C; 71 G; 86 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4.02e-47	Length:	292
Score:	64.00	Matches:	91
Percent Similarity:	97.85%	Conservative:	0
Best Local Similarity:	97.85%	Mismatches:	1
Query Match:	4.48%	Indels:	2
DB:	3	Gaps:	0

US-09-787-097-12 (1-1429) x AAA45505 (1-292)

```
Qy        240 GlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGlu 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         16 GGATTTAATATTACTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAG 75

Qy        260 CysLysIleSerAsnSerSerGlu-ThrValGluCysGluCysSerGluAsnTrpLysGl 279
              ||||||||||||||||||||    ||||||||||||||||||||||||||||||||||||
Db         76 TGTAAGATCAGTAATAGCAGCGA-TACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGG 134

Qy        279 yGluAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCy 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG 194

Qy        299 sAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSe 319
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 CAATTCAAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTC 254

Qy        319 rValProValProAlaAsnGlnSerPheTrpThrArg 331
              ||||||||||||||||||||||||||||||||||||
Db        255 AGTTCCTGTACCAGCTAACCAGTCATTTTGGACTCGA 291
```

RESULT 35
AAI24099/c
ID AAI24099 standard; DNA; 182 BP.
XX
AC AAI24099;
XX
DT 12-OCT-2001 (first entry)
XX
DE Probe #14032 for gene expression analysis in human cervical cell sample.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000670.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human cervical epithelial cells.
XX
PS Claim 25; SEQ ID NO 14032; 487pp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENP). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging of
CC diseases of the cervix, notably cervical cancer. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

```
XX
SQ   Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              8.97e-44        Length:         182
Score:                  60.00           Matches:        60
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            4.20%           Indels:         0
DB:                     4               Gaps:           0

US-09-787-097-12 (1-1429) x AAI24099 (1-182)

Qy        849 SerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGC 123

Qy        869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 TGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGT 63

Qy        889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACC 3

RESULT 36
ABA69223/c
ID   ABA69223 standard; DNA; 182 BP.
XX
AC   ABA69223;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #17528.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000669.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human fetal liver.
XX
PS   Claim 4; SEQ ID NO 17528; 639pp + Sequence Listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for measuring
CC   human gene expression in a sample derived from human foetal liver. The
CC   single exon nucleic acid probes may be used for predicting, measuring and
CC   displaying gene expression in samples derived from human fetal liver. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              8.97e-44        Length:         182
Score:                  60.00           Matches:        60
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            4.20%           Indels:         0
DB:                     4               Gaps:           0

US-09-787-097-12 (1-1429) x ABA69223 (1-182)

Qy        849 SerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGC 123

Qy        869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 TGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGT 63

Qy        889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACC 3

RESULT 37
AAI49392/c
ID   AAI49392 standard; DNA; 182 BP.
XX
AC   AAI49392;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #18078 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human placenta.
XX
PS   Claim 25; SEQ ID NO 18078; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders
XX
SQ   Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;

Alignment Scores:
```

```
Pred. No.:              8.97e-44        Length:         182
Score:                  60.00           Matches:        60
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            4.20%           Indels:         0
DB:                     4               Gaps:           0

US-09-787-097-12 (1-1429) x AAI49392 (1-182)

Qy        849 SerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGC 123

Qy        869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 TGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGT 63

Qy        889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACC 3

RESULT 38
ABA51213/c
ID   ABA51213 standard; DNA; 182 BP.
XX
AC   ABA51213;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human breast cell single exon nucleic acid probe #9908.
XX
KW   Human; microarray; single exon probe; gene expression; breast; disease;
KW   cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157271-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000662.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-496933/54.
XX
PT   New spatially-addressable set of single exon nucleic acid probes, useful
PT   for measuring gene expression in sample derived from human breast,
PT   comprises number of single exon nucleic acid probes.
XX
PS   Claim 4; SEQ ID NO 9908; 327pp + Sequence Listing; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human breast and BT 474 cells. The method involves contacting the
CC   probes with a collection of detectably labelled nucleic acids derived
CC   from mRNA of human breast, and then measuring the label bound to each
CC   probe of the microarray. The probes are useful for verifying the
CC   expression of regions of genomic DNA predicted to encode proteins. They
CC   are useful for gene discovery, and for determining predisposition and/or
CC   prognosing breast disease. Gene expression analysis is useful for
CC   assessing the toxicity of chemical agents on cells. The microarray of
CC   this invention presents a far greater diversity of probes for measuring
CC   gene expression, with far less bias than expressed sequence tag
CC   microarrays. The method is suitable for rapid production of functional
CC   information from genomic sequence. The present sequence is a single exon
CC   nucleic acid probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              8.97e-44        Length:         182
Score:                  60.00           Matches:        60
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            4.20%           Indels:         0
DB:                     4               Gaps:           0

US-09-787-097-12 (1-1429) x ABA51213 (1-182)

Qy        849 SerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGC 123

Qy        869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 TGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGT 63

Qy        889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACC 3

RESULT 39
ABA36140/c
ID   ABA36140 standard; DNA; 182 BP.
XX
AC   ABA36140;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Probe #14606 for gene expression analysis in human heart cell sample.
XX
KW   Human; gene expression; heart; microarray; vascular system; probe;
KW   cardiovascular disease; hypertension; cardiac arrhythmia;
KW   congenital heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157274-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000666.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488899/53.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   hearts.
XX
PS   Claim 4; SEQ ID NO 14606; 530pp; English.
XX
```

CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	8.97e-44	Length:	182
Score:	60.00	Matches:	60
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	4.20%	Indels:	0
DB:	4	Gaps:	0

US-09-787-097-12 (1-1429) x ABA36140 (1-182)

```
Qy        849 SerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGC 123

Qy        869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 TGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGT 63

Qy        889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACC 3
```

RESULT 40
AAK43324/c
ID AAK43324 standard; DNA; 182 BP.
XX
AC AAK43324;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 17881.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000668.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
XX
PS Example 4; SEQ ID NO 17881; 658pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention
XX
SQ Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	8.97e-44	Length:	182
Score:	60.00	Matches:	60
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	4.20%	Indels:	0
DB:	4	Gaps:	0

US-09-787-097-12 (1-1429) x AAK43324 (1-182)

```
Qy        849 SerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGC 123

Qy        869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 TGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGT 63

Qy        889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACC 3
```

RESULT 41
AAK17517/c
ID AAK17517 standard; DNA; 182 BP.
XX
AC AAK17517;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 17508.
XX
KW Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW ss.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000667.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains.
XX

PS Example 4; SEQ ID NO 17508; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention
XX
SQ Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	8.97e-44	Length:	182
Score:	60.00	Matches:	60
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	4.20%	Indels:	0
DB:	4	Gaps:	0

US-09-787-097-12 (1-1429) x AAK17517 (1-182)

```
Qy        849 SerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGC 123

Qy        869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 TGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGT 63

Qy        889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACC 3
```

RESULT 42
ABS42949/c
ID ABS42949 standard; DNA; 182 BP.
XX
AC ABS42949;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human liver single exon probe, SEQ ID No 17939.
XX
KW Human; single exon nucleic acid probe; liver; cirrhosis;
KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW coronary heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157273-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000664.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488898/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.
XX

PS Claim 4; SEQ ID NO 17939; 658pp; English.
XX
CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABS25011-ABS51005 represent human
CC liver single exon nucleic acid probes of the invention. Note: The
CC sequence information for this patent does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	8.97e-44	Length:	182
Score:	60.00	Matches:	60
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	4.20%	Indels:	0
DB:	4	Gaps:	0

US-09-787-097-12 (1-1429) x ABS42949 (1-182)

```
Qy        849 SerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGC 123

Qy        869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 TGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGT 63

Qy        889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACC 3
```

RESULT 43
AAI09674/c
ID AAI09674 standard; DNA; 182 BP.
XX
AC AAI09674;
XX
DT 09-OCT-2001 (first entry)
XX
DE Probe #9665 used to measure gene expression in human breast sample.
XX
KW Probe; human; breast disease; breast cancer; development disorder; ss;
KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS Homo sapiens.
XX
PN WO200157270-A2.
XX
PD 09-AUG-2001.
XX
PF 29-JAN-2001; 2001WO-US000661.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX
DR WPI; 2001-476286/51.
XX
PT Novel single exon nucleic acid probe used to measuring gene expression in
PT a human breast.
XX
PS Claim 25; SEQ ID NO 9665; 322pp; English.
XX
CC The present invention relates to novel single exon nucleic acid probes.
CC The present sequence is one such probe. The probes are useful for
CC measuring human gene expression in a human breast sample, where the probe
CC hybridises at high stringency to a nucleic acid expressed in the human
CC breast. The probes are useful for predicting, diagnosing, grading,
CC staging, monitoring and prognosing diseases of the human breast,
CC particularly those diseases with polygenic aetiology. The diseases
CC include: breast cancer, disorders of development, inflammatory diseases
CC of the breast, fibrocystic changes, proliferative breast disease and non-
CC carcinoma tumours. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	8.97e-44	Length:	182
Score:	60.00	Matches:	60
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	4.20%	Indels:	0
DB:	5	Gaps:	0

US-09-787-097-12 (1-1429) x AAI09674 (1-182)

```
Qy        849 SerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGC 123

Qy        869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 TGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGT 63

Qy        889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACC 3
```

RESULT 44
ABS17404/c
ID ABS17404 standard; DNA; 182 BP.
XX
AC ABS17404;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human genome-derived single exon probe ORF from lung SEQ ID No 17395.
XX
KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesis; pulmonary hypertension;
KW hyaline membrane disease; open reading frame; ORF.
XX
OS Homo sapiens.
XX
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US000665.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2002-114183/15.
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.
XX
PS Claim 4; SEQ ID NO 17395; 634pp; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC ; the novel set of probes which hybridise at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a single exon probe open reading frame of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	8.97e-44	Length:	182
Score:	60.00	Matches:	60
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	4.20%	Indels:	0
DB:	6	Gaps:	0

US-09-787-097-12 (1-1429) x ABS17404 (1-182)

```
Qy        849 SerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGC 123
```

```
Qy        869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 TGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGT 63

Qy        889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACC 3
```

RESULT 45
AAI14910/c
ID AAI14910 standard; DNA; 375 BP.
XX
AC AAI14910;
XX
DT 12-OCT-2001 (first entry)
XX
DE Probe #4843 for gene expression analysis in human cervical cell sample.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000670.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human cervical epithelial cells.
XX
PS Claim 25; SEQ ID NO 4843; 487pp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENP). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging of
CC diseases of the cervix, notably cervical cancer. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	7.92e-41	Length:	375
Score:	57.00	Matches:	57
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.99%	Indels:	0
DB:	4	Gaps:	0

US-09-787-097-12 (1-1429) x AAI14910 (1-375)

```
Qy        865 SerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetPro 884
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 TCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCG 315

Qy        885 SerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeu 904
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 TCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTG 255

Qy        905 LysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro 921
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 AAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCT 204
```

RESULT 46
ABA56648/c
ID ABA56648 standard; DNA; 375 BP.
XX
AC ABA56648;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human foetal liver single exon nucleic acid probe #4953.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human fetal liver.
XX
PS Claim 1; SEQ ID NO 4953; 639pp + Sequence Listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human fetal liver. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	7.92e-41	Length:	375
Score:	57.00	Matches:	57
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.99%	Indels:	0
DB:	4	Gaps:	0

US-09-787-097-12 (1-1429) x ABA56648 (1-375)

```
Qy        865 SerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetPro 884
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        374 TCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCG 315

Qy        885 SerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeu 904
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 TCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTG 255

Qy        905 LysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro 921
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 AAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCT 204
```

RESULT 47
AAI36262/c
ID AAI36262 standard; DNA; 375 BP.
XX
AC AAI36262;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #4948 used to measure gene expression in human placenta sample.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
PS Claim 25; SEQ ID NO 4948; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders
XX
SQ Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	7.92e-41	Length:	375
Score:	57.00	Matches:	57
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.99%	Indels:	0
DB:	4	Gaps:	0

US-09-787-097-12 (1-1429) x AAI36262 (1-375)

```
Qy        865 SerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetPro 884
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 TCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCG 315

Qy        885 SerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeu 904
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 TCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTG 255

Qy        905 LysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro 921
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 AAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCT 204
```

RESULT 48
ABA46100/c
ID ABA46100 standard; DNA; 375 BP.
XX
AC ABA46100;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human breast cell single exon nucleic acid probe #4795.
XX
KW Human; microarray; single exon probe; gene expression; breast; disease;
KW cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157271-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000662.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-496933/54.
XX
PT New spatially-addressable set of single exon nucleic acid probes, useful
PT for measuring gene expression in sample derived from human breast,
PT comprises number of single exon nucleic acid probes.
XX
PS Claim 1; SEQ ID NO 4795; 327pp + Sequence Listing; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting the
CC probes with a collection of detectably labelled nucleic acids derived
CC from mRNA of human breast, and then measuring the label bound to each
CC probe of the microarray. The probes are useful for verifying the
CC expression of regions of genomic DNA predicted to encode proteins. They
CC are useful for gene discovery, and for determining predisposition and/or
CC prognosing breast disease. Gene expression analysis is useful for
CC assessing the toxicity of chemical agents on cells. The microarray of
CC this invention presents a far greater diversity of probes for measuring
CC gene expression, with far less bias than expressed sequence tag
CC microarrays. The method is suitable for rapid production of functional
CC information from genomic sequence. The present sequence is a single exon
CC nucleic acid probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	7.92e-41	Length:	375
Score:	57.00	Matches:	57
Percent Similarity:	100.00%	Conservative:	0

```
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            3.99%         Indels:        0
DB:                     4             Gaps:          0

US-09-787-097-12 (1-1429) x ABA46100 (1-375)

Qy        865 SerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetPro 884
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 TCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCG 315

Qy        885 SerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeu 904
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 TCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTG 255

Qy        905 LysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro 921
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 AAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCT 204

RESULT 49
ABA26254/c
ID   ABA26254 standard; DNA; 375 BP.
XX
AC   ABA26254;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Probe #4720 for gene expression analysis in human heart cell sample.
XX
KW   Human; gene expression; heart; microarray; vascular system; probe;
KW   cardiovascular disease; hypertension; cardiac arrhythmia;
KW   congenital heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157274-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000666.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488899/53.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   hearts.
XX
PS   Claim 1; SEQ ID NO 4720; 530pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes for
CC   measuring human gene expression in a sample derived from human heart. The
CC   present sequence is one such probe. The probes may be used for
CC   predicting, measuring and displaying gene expression in samples derived
CC   from the human heart via microarrays. By measuring gene expression, the
CC   probes are useful for predicting, diagnosing, grading, staging,
CC   monitoring and prognosing diseases of the human heart and vascular system
CC   e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC   congenital heart disease. Note: The sequence data for this patent did not
CC   form part of the printed specification, but was obtained in electronic
CC   format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              7.92e-41      Length:        375
Score:                  57.00         Matches:       57
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            3.99%         Indels:        0
DB:                     4             Gaps:          0

US-09-787-097-12 (1-1429) x ABA26254 (1-375)

Qy        865 SerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetPro 884
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 TCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCG 315

Qy        885 SerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeu 904
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 TCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTG 255

Qy        905 LysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro 921
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 AAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCT 204

RESULT 50
AAK30307/c
ID   AAK30307 standard; DNA; 375 BP.
XX
AC   AAK30307;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human bone marrow expressed single exon probe SEQ ID NO: 4864.
XX
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157276-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000668.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human bone marrow.
XX
PS   Example 4; SEQ ID NO 4864; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention
XX
SQ   Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              7.92e-41      Length:        375
```

```
Score:                  57.00          Matches:        57
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            3.99%          Indels:         0
DB:                     4              Gaps:           0

US-09-787-097-12 (1-1429) x AAK30307 (1-375)

Qy        865 SerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetPro 884
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 TCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCG 315

Qy        885 SerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeu 904
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 TCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTG 255

Qy        905 LysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro 921
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 AAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCT 204

RESULT 51
AAK04785/c
ID   AAK04785 standard; DNA; 375 BP.
XX
AC   AAK04785;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human brain expressed single exon probe SEQ ID NO: 4776.
XX
KW   Human; brain expressed exon; gene expression analysis; probe; microarray;
KW   Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW   ss.
XX
OS   Homo sapiens.
XX
PN   WO200157275-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000667.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   brains.
XX
PS   Example 4; SEQ ID NO 4776; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is one of the probes of the
CC   invention
XX
SQ   Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              7.92e-41       Length:         375
```

```
Score:                  57.00          Matches:        57
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            3.99%          Indels:         0
DB:                     4              Gaps:           0

US-09-787-097-12 (1-1429) x AAK04785 (1-375)

Qy        865 SerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetPro 884
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 TCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCG 315

Qy        885 SerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeu 904
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 TCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTG 255

Qy        905 LysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro 921
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 AAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCT 204

RESULT 52
ABS29950/c
ID   ABS29950 standard; DNA; 375 BP.
XX
AC   ABS29950;
XX
DT   25-FEB-2003  (first entry)
XX
DE   Human liver single exon probe, SEQ ID No 4940.
XX
KW   Human; single exon nucleic acid probe; liver; cirrhosis;
KW   hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW   coronary heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157273-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000664.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488898/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human adult liver.
XX
PS   Claim 1; SEQ ID NO 4940; 658pp; English.
XX
CC   The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC   measuring human gene expression in a sample derived from human adult
CC   liver, comprising one of 13109 defined nucleotide sequences given in the
CC   specification (or complements/ fragments). The probe hybridises at high
CC   stringency to a nucleic acid molecule expressed in the human adult liver.
CC   (I) may be used for predicting, measuring and displaying gene expression
CC   in samples derived from human adult liver. The genes identified may be
CC   involved in genetic liver diseases such as cirrhosis,
CC   hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC   associated with coronary heart disease. ABS25011-ABS51005 represent human
CC   liver single exon nucleic acid probes of the invention. Note: The
CC   sequence information for this patent does not appear in the printed
```

```
CC   specification but was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              7.92e-41      Length:         375
Score:                  57.00         Matches:        57
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            3.99%         Indels:         0
DB:                     4             Gaps:           0

US-09-787-097-12 (1-1429) x ABS29950 (1-375)

Qy        865 SerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetPro 884
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 TCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCG 315

Qy        885 SerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeu 904
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 TCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTG 255

Qy        905 LysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro 921
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 AAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCT 204

RESULT 53
AAI04689/c
ID   AAI04689 standard; DNA; 375 BP.
XX
AC   AAI04689;
XX
DT   09-OCT-2001  (first entry)
XX
DE   Probe #4680 used to measure gene expression in human breast sample.
XX
KW   Probe; human; breast disease; breast cancer; development disorder; ss;
KW   inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS   Homo sapiens.
XX
PN   WO200157270-A2.
XX
PD   09-AUG-2001.
XX
PF   29-JAN-2001; 2001WO-US000661.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-476286/51.
XX
PT   Novel single exon nucleic acid probe used to measuring gene expression in
PT   a human breast.
XX
PS   Claim 25; SEQ ID NO 4680; 322pp; English.
XX
CC   The present invention relates to novel single exon nucleic acid probes.
CC   The present sequence is one such probe. The probes are useful for
CC   measuring human gene expression in a human breast sample, where the probe
CC   hybridises at high stringency to a nucleic acid expressed in the human
CC   breast. The probes are useful for predicting, diagnosing, grading,
CC   staging, monitoring and prognosing diseases of the human breast,
CC   particularly those diseases with polygenic aetiology. The diseases
CC   include: breast cancer, disorders of development, inflammatory diseases
CC   of the breast, fibrocystic changes, proliferative breast disease and non-
CC   carcinoma tumours. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              7.92e-41      Length:         375
Score:                  57.00         Matches:        57
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            3.99%         Indels:         0
DB:                     5             Gaps:           0

US-09-787-097-12 (1-1429) x AAI04689 (1-375)

Qy        865 SerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetPro 884
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 TCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCG 315

Qy        885 SerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeu 904
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 TCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTG 255

Qy        905 LysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro 921
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 AAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCT 204

RESULT 54
ABS04892/c
ID   ABS04892 standard; DNA; 375 BP.
XX
AC   ABS04892;
XX
DT   19-AUG-2002  (first entry)
XX
DE   Human genome-derived single exon probe from lung SEQ ID No 4883.
XX
KW   Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease.
XX
OS   Homo sapiens.
XX
PN   WO200186003-A2.
XX
PD   15-NOV-2001.
XX
PF   30-JAN-2001; 2001WO-US000665.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2002-114183/15.
XX
```

```
PT   Spatially-addressable set of single exon nucleic acid probes, used to
PT   measure gene expression in human lung samples.
XX
PS   Claim 1; SEQ ID NO 4883; 634pp; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human lung comprising single exon nucleic acid probes having one of
CC   12614 nucleic acid sequences mentioned in the specification, or their
CC   complements or the 12387 open reading frames derived from the 12614
CC   probes. Also included are a microarray comprising the novel set of probes
CC   ; the novel set of probes which hybridise at high stringency to a nucleic
CC   acid expressed in the human lung; measuring gene expression in a sample
CC   derived from human lung, comprising (a) contacting the array with a
CC   collection of detectably labeled nucleic acids derived from human lung
CC   mRNA, and (b) measuring the label detectably bound to each probe of the
CC   array; identifying exons in a eukaryotic genome, comprising (a)
CC   algorithmically predicting at least one exon from genomic sequences of
CC   the eukaryote; and (b) detecting specific hybridisation of detectably
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC   having a fragment identical to the predicted exon, the probe is included
CC   in the above mentioned microarray; assigning exons to a single gene,
CC   comprising (a) identifying exons from genomic sequence by the method
CC   above and (b) measuring the expression of each of the exons in several
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern of
CC   expression of the exons in the tissues and/or cell types indicates that
CC   the exons should be assigned to a single gene; a peptide comprising one
CC   of 12011 sequences, mentioned in the specification, or encoded by the
CC   probes/open reading frames (ORF). The probes are used for gene expression
CC   analysis, and for identifying exons in a gene, particularly using human
CC   lung derived mRNA and for the study of lung diseases such as asthma, lung
CC   cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC   disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC   tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC   Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC   histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC   Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC   dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC   present sequence is a single exon probe of the invention. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                7.92e-41        Length:          375
Score:                    57.00           Matches:         57
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              3.99%           Indels:          0
DB:                       6               Gaps:            0

US-09-787-097-12 (1-1429) x ABS04892 (1-375)

Qy        865 SerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetPro 884
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 TCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCG 315

Qy        885 SerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeu 904
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 TCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTG 255

Qy        905 LysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro 921
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 AAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCT 204

RESULT 55
AAZ91926/c
ID   AAZ91926 standard; DNA; 226 BP.
XX
AC   AAZ91926;
XX
DT   08-JUN-2000  (first entry)
XX
DE   Wild type murine mahogany gene exon 15 sequence.
XX
KW   Mahogany gene; mouse; mg gene; regulatory defect; gene therapy; obesity;
KW   weight regulation; cell therapy; body weight disorder; cachexia;
KW   anorexia; hyperpigmentation; increased metabolic rate disorder;
KW   hyperphagia; Antiobesity; antianorexic; anticachexic; ds.
XX
OS   Mus sp.
XX
PN   WO200005373-A2.
XX
PD   03-FEB-2000.
XX
PF   21-JUL-1999;   99WO-US016484.
XX
PR   21-JUL-1998;   98US-0093630P.
PR   20-OCT-1998;   98US-0104978P.
PR   05-FEB-1999;   99US-00245041.
XX
PA   (MILL-) MILLENIUM PHARM INC.
XX
PI   Moore K, Nagle DL;
XX
DR   WPI; 2000-195103/17.
XX
PT   New human and murine mahogany genes, useful, e.g. for diagnosis and
PT   treatment of body weight disorders.
XX
PS   Example; Fig 6; 188pp; English.
XX
CC   This sequence is a fragment of a murine mahogany gene of the invention.
CC   The mahogany genes are used: (i) to produce recombinant mahogany (mg)
CC   proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming
CC   therapeutics; (iii) as a source of diagnostic probes and primers for
CC   detecting expression of mg genes or mutations, regulatory defects, in
CC   this gene, or for isolation of related sequences; and (iv) in (cell-
CC   based) gene therapy. (II) are used to raise specific antibodies (Ab); to
CC   identify other (extra)cellular products involved in weight regulation,
CC   and to screen for agents that disrupt interaction between (II) and other
CC   macromolecules. The Ab are used to detect abnormal levels (or function)
CC   of (II) (for diagnosis, prognosis or monitoring of treatment); to
CC   evaluate (II)-expressing cells intended for cell therapy, and as
CC   therapeutic mg inhibitors. Cells that express the mg gene (or contain the
CC   mg polypeptide) are used to identify agents (A) that modulate mg
CC   activity. (A) are potentially useful for the treatment of body weight
CC   disorders, particularly obesity, cachexia or anorexia, or other
CC   conditions associated with the mg gene such as hyperpigmentation,
CC   hyperphagia and disorders that result in increased metabolic rate
XX
SQ   Sequence 226 BP; 51 A; 66 C; 52 G; 57 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.01e-35        Length:          226
Score:                    51.00           Matches:         51
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              3.57%           Indels:          0
DB:                       3               Gaps:            0

US-09-787-097-12 (1-1429) x AAZ91926 (1-226)

Qy        871 AspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSerAspAla 890
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        225 GATATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGTGATGCT 166

Qy        891 GlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThrCysIle 910
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        165 GGCTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGATTAAAGGCTGCAACCTGCATC 106

Qy        911 AsnProLeuAsnGlySerValCysGluArgPro 921
```

```
                    ||||||||||||||||||||||||||||||||||
Db        105 AACCCTCTCAATGGCAGCGTCTGTGAAAGGCCT 73

RESULT 56
AAS69062
ID   AAS69062 standard; cDNA; 2267 BP.
XX
AC   AAS69062;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #4866.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG04875.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 4866; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 2267 BP; 917 A; 455 C; 453 G; 442 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              5.04e-33        Length:         2267
Score:                  49.00           Matches:        49
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            3.43%           Indels:         0
DB:                     5               Gaps:           0

US-09-787-097-12 (1-1429) x AAS69062 (1-2267)

Qy       1140 CysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeu 1159
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        163 TGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTT 222

Qy       1160 LeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAla 1179
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 CTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCT 282

Qy       1180 IleAsnPheValAlaThrProAspGlu 1188
                |||||||||||||||||||||||||||
Db        283 ATCAATTTTGTGGCTACTCCTGACGAA 309

RESULT 57
AAI13789/c
ID   AAI13789 standard; DNA; 402 BP.
XX
AC   AAI13789;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #3722 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000670.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human cervical epithelial cells.
XX
PS   Claim 25; SEQ ID NO 3722; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging of
CC   diseases of the cervix, notably cervical cancer. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              4.58e-31        Length:         402
Score:                  46.00           Matches:        46
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            3.22%           Indels:         0
```

```
DB:                    4             Gaps:          0

US-09-787-097-12 (1-1429) x AAI13789 (1-402)

Qy       1062 AlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGlu 1081
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 GCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAG 330

Qy       1082 AsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThr 1101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 AACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACC 270

Qy       1102 AsnGlyGlyLysCysGln 1107
              |||||||||||||||||
Db        269 AATGGAGGGAAATGTCAG 252

RESULT 58
ABA55499/c
ID   ABA55499 standard; DNA; 402 BP.
XX
AC   ABA55499;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #3804.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000669.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human fetal liver.
XX
PS   Claim 1; SEQ ID NO 3804; 639pp + Sequence Listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for measuring
CC   human gene expression in a sample derived from human foetal liver. The
CC   single exon nucleic acid probes may be used for predicting, measuring and
CC   displaying gene expression in samples derived from human fetal liver. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              4.58e-31        Length:         402
Score:                  46.00           Matches:        46
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            3.22%           Indels:         0
DB:                     4               Gaps:           0

US-09-787-097-12 (1-1429) x ABA55499 (1-402)

Qy       1062 AlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGlu 1081
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 GCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAG 330

Qy       1082 AsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThr 1101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 AACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACC 270

Qy       1102 AsnGlyGlyLysCysGln 1107
              |||||||||||||||||
Db        269 AATGGAGGGAAATGTCAG 252

RESULT 59
AAI35149/c
ID   AAI35149 standard; DNA; 402 BP.
XX
AC   AAI35149;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #3835 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human placenta.
XX
PS   Claim 25; SEQ ID NO 3835; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders
XX
SQ   Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              4.58e-31        Length:         402
Score:                  46.00           Matches:        46
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            3.22%           Indels:         0
DB:                     4               Gaps:           0

US-09-787-097-12 (1-1429) x AAI35149 (1-402)
```

```
Qy       1062 AlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGlu 1081
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 GCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAG 330

Qy       1082 AsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThr 1101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 AACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACC 270

Qy       1102 AsnGlyGlyLysCysGln 1107
              ||||||||||||||||||
Db        269 AATGGAGGGAAATGTCAG 252

RESULT 60
ABA45027/c
ID   ABA45027 standard; DNA; 402 BP.
XX
AC   ABA45027;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human breast cell single exon nucleic acid probe #3722.
XX
KW   Human; microarray; single exon probe; gene expression; breast; disease;
KW   cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157271-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000662.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-496933/54.
XX
PT   New spatially-addressable set of single exon nucleic acid probes, useful
PT   for measuring gene expression in sample derived from human breast,
PT   comprises number of single exon nucleic acid probes.
XX
PS   Claim 1; SEQ ID NO 3722; 327pp + Sequence Listing; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human breast and BT 474 cells. The method involves contacting the
CC   probes with a collection of detectably labelled nucleic acids derived
CC   from mRNA of human breast, and then measuring the label bound to each
CC   probe of the microarray. The probes are useful for verifying the
CC   expression of regions of genomic DNA predicted to encode proteins. They
CC   are useful for gene discovery, and for determining predisposition and/or
CC   prognosing breast disease. Gene expression analysis is useful for
CC   assessing the toxicity of chemical agents on cells. The microarray of
CC   this invention presents a far greater diversity of probes for measuring
CC   gene expression, with far less bias than expressed sequence tag
CC   microarrays. The method is suitable for rapid production of functional
CC   information from genomic sequence. The present sequence is a single exon
CC   nucleic acid probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                4.58e-31        Length:          402
Score:                    46.00           Matches:         46
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              3.22%           Indels:          0
DB:                       4               Gaps:            0

US-09-787-097-12 (1-1429) x ABA45027 (1-402)

Qy       1062 AlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGlu 1081
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 GCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAG 330

Qy       1082 AsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThr 1101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 AACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACC 270

Qy       1102 AsnGlyGlyLysCysGln 1107
              ||||||||||||||||||
Db        269 AATGGAGGGAAATGTCAG 252

RESULT 61
ABA25215/c
ID   ABA25215 standard; DNA; 402 BP.
XX
AC   ABA25215;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Probe #3681 for gene expression analysis in human heart cell sample.
XX
KW   Human; gene expression; heart; microarray; vascular system; probe;
KW   cardiovascular disease; hypertension; cardiac arrhythmia;
KW   congenital heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157274-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000666.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488899/53.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   hearts.
XX
PS   Claim 1; SEQ ID NO 3681; 530pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes for
CC   measuring human gene expression in a sample derived from human heart. The
CC   present sequence is one such probe. The probes may be used for
CC   predicting, measuring and displaying gene expression in samples derived
CC   from the human heart via microarrays. By measuring gene expression, the
CC   probes are useful for predicting, diagnosing, grading, staging,
CC   monitoring and prognosing diseases of the human heart and vascular system
CC   e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC   congenital heart disease. Note: The sequence data for this patent did not
```

```
CC   form part of the printed specification, but was obtained in electronic
CC   format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              4.58e-31        Length:         402
Score:                  46.00           Matches:        46
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            3.22%           Indels:         0
DB:                     4               Gaps:           0

US-09-787-097-12 (1-1429) x ABA25215 (1-402)

Qy       1062 AlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGlu 1081
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 GCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAG 330

Qy       1082 AsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThr 1101
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 AACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACC 270

Qy       1102 AsnGlyGlyLysCysGln 1107
             |||||||||||||||||
Db        269 AATGGAGGGAAATGTCAG 252

RESULT 62
AAK29197/c
ID   AAK29197 standard; DNA; 402 BP.
XX
AC   AAK29197;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human bone marrow expressed single exon probe SEQ ID NO: 3754.
XX
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157276-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000668.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human bone marrow.
XX
PS   Example 4; SEQ ID NO 3754; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention
XX
SQ   Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              4.58e-31        Length:         402
Score:                  46.00           Matches:        46
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            3.22%           Indels:         0
DB:                     4               Gaps:           0

US-09-787-097-12 (1-1429) x AAK29197 (1-402)

Qy       1062 AlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGlu 1081
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 GCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAG 330

Qy       1082 AsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThr 1101
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 AACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACC 270

Qy       1102 AsnGlyGlyLysCysGln 1107
             |||||||||||||||||
Db        269 AATGGAGGGAAATGTCAG 252

RESULT 63
AAK03733/c
ID   AAK03733 standard; DNA; 402 BP.
XX
AC   AAK03733;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human brain expressed single exon probe SEQ ID NO: 3724.
XX
KW   Human; brain expressed exon; gene expression analysis; probe; microarray;
KW   Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW   ss.
XX
OS   Homo sapiens.
XX
PN   WO200157275-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000667.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   brains.
XX
PS   Example 4; SEQ ID NO 3724; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is one of the probes of the
CC   invention
```

```
XX
SQ   Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              4.58e-31        Length:         402
Score:                  46.00           Matches:        46
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            3.22%           Indels:         0
DB:                     4               Gaps:           0

US-09-787-097-12 (1-1429) x AAK03733 (1-402)

Qy       1062 AlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGlu 1081
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 GCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAG 330

Qy       1082 AsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThr 1101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 AACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACC 270

Qy       1102 AsnGlyGlyLysCysGln 1107
              ||||||||||||||||||
Db        269 AATGGAGGGAAATGTCAG 252

RESULT 64
ABS28814/c
ID   ABS28814 standard; DNA; 402 BP.
XX
AC   ABS28814;
XX
DT   25-FEB-2003  (first entry)
XX
DE   Human liver single exon probe, SEQ ID No 3804.
XX
KW   Human; single exon nucleic acid probe; liver; cirrhosis;
KW   hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW   coronary heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157273-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000664.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488898/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human adult liver.
XX
PS   Claim 1; SEQ ID NO 3804; 658pp; English.
XX
CC   The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC   measuring human gene expression in a sample derived from human adult
CC   liver, comprising one of 13109 defined nucleotide sequences given in the
CC   specification (or complements/ fragments). The probe hybridises at high
CC   stringency to a nucleic acid molecule expressed in the human adult liver.
CC   (I) may be used for predicting, measuring and displaying gene expression
CC   in samples derived from human adult liver. The genes identified may be
CC   involved in genetic liver diseases such as cirrhosis,
CC   hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC   associated with coronary heart disease. ABS25011-ABS51005 represent human
CC   liver single exon nucleic acid probes of the invention. Note: The
CC   sequence information for this patent does not appear in the printed
CC   specification but was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              4.58e-31        Length:         402
Score:                  46.00           Matches:        46
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            3.22%           Indels:         0
DB:                     4               Gaps:           0

US-09-787-097-12 (1-1429) x ABS28814 (1-402)

Qy       1062 AlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGlu 1081
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 GCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAG 330

Qy       1082 AsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThr 1101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 AACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACC 270

Qy       1102 AsnGlyGlyLysCysGln 1107
              ||||||||||||||||||
Db        269 AATGGAGGGAAATGTCAG 252

RESULT 65
AAI03665/c
ID   AAI03665 standard; DNA; 402 BP.
XX
AC   AAI03665;
XX
DT   09-OCT-2001  (first entry)
XX
DE   Probe #3656 used to measure gene expression in human breast sample.
XX
KW   Probe; human; breast disease; breast cancer; development disorder; ss;
KW   inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS   Homo sapiens.
XX
PN   WO200157270-A2.
XX
PD   09-AUG-2001.
XX
PF   29-JAN-2001; 2001WO-US000661.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-476286/51.
XX
PT   Novel single exon nucleic acid probe used to measuring gene expression in
PT   a human breast.
XX
PS   Claim 25; SEQ ID NO 3656; 322pp; English.
XX
CC   The present invention relates to novel single exon nucleic acid probes.
```

CC The present sequence is one such probe. The probes are useful for
CC measuring human gene expression in a human breast sample, where the probe
CC hybridises at high stringency to a nucleic acid expressed in the human
CC breast. The probes are useful for predicting, diagnosing, grading,
CC staging, monitoring and prognosing diseases of the human breast,
CC particularly those diseases with polygenic aetiology. The diseases
CC include: breast cancer, disorders of development, inflammatory diseases
CC of the breast, fibrocystic changes, proliferative breast disease and non-
CC carcinoma tumours. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4.58e-31	Length:	402
Score:	46.00	Matches:	46
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.22%	Indels:	0
DB:	5	Gaps:	0

US-09-787-097-12 (1-1429) x AAI03665 (1-402)

```
Qy       1062 AlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGlu 1081
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 GCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAG 330

Qy       1082 AsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThr 1101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 AACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACC 270

Qy       1102 AsnGlyGlyLysCysGln 1107
              |||||||||||||||||
Db        269 AATGGAGGGAAATGTCAG 252
```

RESULT 66
ABS03749/c
ID ABS03749 standard; DNA; 402 BP.
XX
AC ABS03749;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human genome-derived single exon probe from lung SEQ ID No 3740.
XX
KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesis; pulmonary hypertension;
KW hyaline membrane disease.
XX
OS Homo sapiens.
XX
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US000665.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2002-114183/15.
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.
XX
PS Claim 1; SEQ ID NO 3740; 634pp; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC ; the novel set of probes which hybridise at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a single exon probe of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4.58e-31	Length:	402
Score:	46.00	Matches:	46
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.22%	Indels:	0
DB:	6	Gaps:	0

US-09-787-097-12 (1-1429) x ABS03749 (1-402)

```
Qy       1062 AlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGlu 1081
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 GCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAG 330

Qy       1082 AsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThr 1101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 AACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACC 270

Qy       1102 AsnGlyGlyLysCysGln 1107
              |||||||||||||||||
Db        269 AATGGAGGGAAATGTCAG 252
```

RESULT 67
AAI23009/c
ID AAI23009 standard; DNA; 137 BP.
XX
AC AAI23009;
XX
DT 12-OCT-2001 (first entry)
XX
DE Probe #12942 for gene expression analysis in human cervical cell sample.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000670.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human cervical epithelial cells.
XX
PS Claim 25; SEQ ID NO 12942; 487pp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENP). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging of
CC diseases of the cervix, notably cervical cancer. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.29e-30	Length:	137
Score:	45.00	Matches:	45
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.15%	Indels:	0
DB:	4	Gaps:	0

US-09-787-097-12 (1-1429) x AAI23009 (1-137)

```
Qy       1063 CysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsn 1082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 TGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAAC 76

Qy       1083 LeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsn 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          75 CTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAAT 16

Qy       1103 GlyGlyLysCysGln 1107
              ||||||||||||||||
Db          15 GGAGGGAAATGTCAG 1
```

RESULT 68
ABA68099/c
ID ABA68099 standard; DNA; 137 BP.
XX
AC ABA68099;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human foetal liver single exon nucleic acid probe #16404.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human fetal liver.
XX
PS Claim 4; SEQ ID NO 16404; 639pp + Sequence Listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human fetal liver. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.29e-30	Length:	137
Score:	45.00	Matches:	45
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.15%	Indels:	0
DB:	4	Gaps:	0

US-09-787-097-12 (1-1429) x ABA68099 (1-137)

```
Qy       1063 CysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsn 1082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 TGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAAC 76

Qy       1083 LeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsn 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          75 CTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAAT 16

Qy       1103 GlyGlyLysCysGln 1107
              ||||||||||||||||
Db          15 GGAGGGAAATGTCAG 1
```

```
RESULT 69
AAI48309/c
ID   AAI48309 standard; DNA; 137 BP.
XX
AC   AAI48309;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #16995 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human placenta.
XX
PS   Claim 25; SEQ ID NO 16995; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders
XX
SQ   Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.29e-30      Length:          137
Score:                  45.00         Matches:         45
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            3.15%         Indels:          0
DB:                     4             Gaps:            0

US-09-787-097-12 (1-1429) x AAI48309 (1-137)

Qy       1063 CysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsn 1082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 TGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAAC 76

Qy       1083 LeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsn 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          75 CTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAAT 16

Qy       1103 GlyGlyLysCysGln 1107
              |||||||||||||||
Db          15 GGAGGGAAATGTCAG 1

RESULT 70
ABA50165/c
ID   ABA50165 standard; DNA; 137 BP.
XX
AC   ABA50165;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human breast cell single exon nucleic acid probe #8860.
XX
KW   Human; microarray; single exon probe; gene expression; breast; disease;
KW   cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157271-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000662.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-496933/54.
XX
PT   New spatially-addressable set of single exon nucleic acid probes, useful
PT   for measuring gene expression in sample derived from human breast,
PT   comprises number of single exon nucleic acid probes.
XX
PS   Claim 4; SEQ ID NO 8860; 327pp + Sequence Listing; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human breast and BT 474 cells. The method involves contacting the
CC   probes with a collection of detectably labelled nucleic acids derived
CC   from mRNA of human breast, and then measuring the label bound to each
CC   probe of the microarray. The probes are useful for verifying the
CC   expression of regions of genomic DNA predicted to encode proteins. They
CC   are useful for gene discovery, and for determining predisposition and/or
CC   prognosing breast disease. Gene expression analysis is useful for
CC   assessing the toxicity of chemical agents on cells. The microarray of
CC   this invention presents a far greater diversity of probes for measuring
CC   gene expression, with far less bias than expressed sequence tag
CC   microarrays. The method is suitable for rapid production of functional
CC   information from genomic sequence. The present sequence is a single exon
CC   nucleic acid probe of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.29e-30      Length:          137
Score:                  45.00         Matches:         45
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            3.15%         Indels:          0
DB:                     4             Gaps:            0

US-09-787-097-12 (1-1429) x ABA50165 (1-137)

Qy       1063 CysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsn 1082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 TGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAAC 76
```

```
Qy       1083 LeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsn 1102
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAAT 16

Qy       1103 GlyGlyLysCysGln 1107
             ||||||||||||||||
Db         15 GGAGGGAAATGTCAG 1
```

RESULT 71
ABA35129/c
ID ABA35129 standard; DNA; 137 BP.
XX
AC ABA35129;
XX
DT 23-JAN-2002 (first entry)
XX
DE Probe #13595 for gene expression analysis in human heart cell sample.
XX
KW Human; gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000666.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
XX
PS Claim 4; SEQ ID NO 13595; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.29e-30	Length:	137
Score:	45.00	Matches:	45
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.15%	Indels:	0
DB:	4	Gaps:	0

US-09-787-097-12 (1-1429) x ABA35129 (1-137)

```
Qy       1063 CysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsn 1082
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 TGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAAC 76

Qy       1083 LeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsn 1102
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAAT 16

Qy       1103 GlyGlyLysCysGln 1107
             ||||||||||||||||
Db         15 GGAGGGAAATGTCAG 1
```

RESULT 72
AAK42239/c
ID AAK42239 standard; DNA; 137 BP.
XX
AC AAK42239;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 16796.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000668.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
XX
PS Example 4; SEQ ID NO 16796; 658pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention
XX
SQ Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.29e-30	Length:	137
Score:	45.00	Matches:	45
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.15%	Indels:	0
DB:	4	Gaps:	0

US-09-787-097-12 (1-1429) x AAK42239 (1-137)

Qy 1063 CysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsn 1082

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 TGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAAC 76

Qy       1083 LeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsn 1102
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAAT 16

Qy       1103 GlyGlyLysCysGln 1107
                ||||||||||||||||
Db         15 GGAGGGAAATGTCAG 1

RESULT 73
AAK16486/c
ID   AAK16486 standard; DNA; 137 BP.
XX
AC   AAK16486;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human brain expressed single exon probe SEQ ID NO: 16477.
XX
KW   Human; brain expressed exon; gene expression analysis; probe; microarray;
KW   Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW   ss.
XX
OS   Homo sapiens.
XX
PN   WO200157275-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000667.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   brains.
XX
PS   Example 4; SEQ ID NO 16477; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is one of the probes of the
CC   invention
XX
SQ   Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.29e-30        Length:         137
Score:                  45.00           Matches:        45
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            3.15%           Indels:         0
DB:                     4               Gaps:           0

US-09-787-097-12 (1-1429) x AAK16486 (1-137)

Qy       1063 CysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsn 1082
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 TGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAAC 76

Qy       1083 LeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsn 1102
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAAT 16

Qy       1103 GlyGlyLysCysGln 1107
                ||||||||||||||||
Db         15 GGAGGGAAATGTCAG 1

RESULT 74
ABS41845/c
ID   ABS41845 standard; DNA; 137 BP.
XX
AC   ABS41845;
XX
DT   25-FEB-2003  (first entry)
XX
DE   Human liver single exon probe, SEQ ID No 16835.
XX
KW   Human; single exon nucleic acid probe; liver; cirrhosis;
KW   hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW   coronary heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157273-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000664.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488898/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human adult liver.
XX
PS   Claim 4; SEQ ID NO 16835; 658pp; English.
XX
CC   The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC   measuring human gene expression in a sample derived from human adult
CC   liver, comprising one of 13109 defined nucleotide sequences given in the
CC   specification (or complements/ fragments). The probe hybridises at high
CC   stringency to a nucleic acid molecule expressed in the human adult liver.
CC   (I) may be used for predicting, measuring and displaying gene expression
CC   in samples derived from human adult liver. The genes identified may be
CC   involved in genetic liver diseases such as cirrhosis,
CC   hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC   associated with coronary heart disease. ABS25011-ABS51005 represent human
CC   liver single exon nucleic acid probes of the invention. Note: The
CC   sequence information for this patent does not appear in the printed
CC   specification but was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.29e-30        Length:         137
Score:                  45.00           Matches:        45
Percent Similarity:     100.00%         Conservative:   0
```

```
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            3.15%          Indels:        0
DB:                     4              Gaps:          0

US-09-787-097-12 (1-1429) x ABS41845 (1-137)

Qy       1063 CysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsn 1082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 TGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAAC 76

Qy       1083 LeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsn 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          75 CTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAAT 16

Qy       1103 GlyGlyLysCysGln 1107
              ||||||||||||||||
Db          15 GGAGGGAAATGTCAG 1

RESULT 75
AAI08671/c
ID   AAI08671 standard; DNA; 137 BP.
XX
AC   AAI08671;
XX
DT   09-OCT-2001  (first entry)
XX
DE   Probe #8662 used to measure gene expression in human breast sample.
XX
KW   Probe; human; breast disease; breast cancer; development disorder; ss;
KW   inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS   Homo sapiens.
XX
PN   WO200157270-A2.
XX
PD   09-AUG-2001.
XX
PF   29-JAN-2001; 2001WO-US000661.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-476286/51.
XX
PT   Novel single exon nucleic acid probe used to measuring gene expression in
PT   a human breast.
XX
PS   Claim 25; SEQ ID NO 8662; 322pp; English.
XX
CC   The present invention relates to novel single exon nucleic acid probes.
CC   The present sequence is one such probe. The probes are useful for
CC   measuring human gene expression in a human breast sample, where the probe
CC   hybridises at high stringency to a nucleic acid expressed in the human
CC   breast. The probes are useful for predicting, diagnosing, grading,
CC   staging, monitoring and prognosing diseases of the human breast,
CC   particularly those diseases with polygenic aetiology. The diseases
CC   include: breast cancer, disorders of development, inflammatory diseases
CC   of the breast, fibrocystic changes, proliferative breast disease and non-
CC   carcinoma tumours. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.29e-30       Length:        137
Score:                  45.00          Matches:       45
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            3.15%          Indels:        0
DB:                     5              Gaps:          0

US-09-787-097-12 (1-1429) x AAI08671 (1-137)

Qy       1063 CysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsn 1082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 TGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAAC 76

Qy       1083 LeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsn 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          75 CTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAAT 16

Qy       1103 GlyGlyLysCysGln 1107
              ||||||||||||||||
Db          15 GGAGGGAAATGTCAG 1


Search completed: March  1, 2004, 07:55:33
Job time : 1707 secs
```

This page blank (usptо)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 1, 2004, 07:35:35 ; Search time 1026 Seconds
(without alignments)
5024.435 Million cell updates/sec

Title: US-09-787-097-12
Perfect score: 1429
Sequence: 1 MVAAAAATEARLRRRTAATA.....KSGAVRNRKQQPPAQPGTCI 1429

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2353733 seqs, 1803733377 residues

Word size: 1

Total number of hits satisfying chosen parameters: 4700994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09787097/runat_26022004_063017_23865/app_query.fasta_1.1607
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=oligo.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=100 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1
-ALIGN=75 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09787097_@CGN_1_1_136_@runat_26022004_063017_23865
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1141	79.8	8589	9	US-09-893-238-14	Sequence 14, Appl
	2	979	68.5	4072	9	US-09-893-238-16	Sequence 16, Appl
	3	503	35.2	6370	9	US-09-893-238-12	Sequence 12, Appl
	4	496	34.7	2625	9	US-09-893-238-18	Sequence 18, Appl
	5	206	14.4	8827	9	US-09-893-238-1	Sequence 1, Appli
	6	94	6.6	2419	9	US-09-893-238-8	Sequence 8, Appli
c	7	92	6.4	625	14	US-10-198-846-5869	Sequence 5869, Ap
c	8	82	5.7	246	9	US-09-736-457-966	Sequence 966, App
c	9	82	5.7	246	9	US-09-902-941-966	Sequence 966, App
c	10	82	5.7	246	9	US-09-849-626-966	Sequence 966, App
c	11	82	5.7	246	14	US-10-017-754-966	Sequence 966, App
c	12	82	5.7	246	14	US-10-113-872-966	Sequence 966, App
	13	77	5.4	207433	15	US-10-277-216-5	Sequence 5, Appli
	14	77	5.4	207433	16	US-10-126-022-5	Sequence 5, Appli
	15	76	5.3	1051	9	US-09-893-238-10	Sequence 10, Appl
	16	73	5.1	90050	9	US-09-893-238-5	Sequence 5, Appli
	17	69	4.8	531	9	US-09-864-761-15900	Sequence 15900, A
	18	64	4.5	193	9	US-09-864-761-32404	Sequence 32404, A
c	19	60	4.2	182	9	US-09-864-761-21460	Sequence 21460, A
c	20	57	4.0	375	9	US-09-864-761-4720	Sequence 4720, Ap
c	21	51	3.6	226	9	US-09-893-238-6	Sequence 6, Appli
	22	46	3.2	212	9	US-09-796-692-7610	Sequence 7610, Ap
	23	46	3.2	212	14	US-10-040-862-7610	Sequence 7610, Ap
	24	46	3.2	212	15	US-10-057-475B-7610	Sequence 7610, Ap
	25	46	3.2	212	15	US-10-154-884B-7610	Sequence 7610, Ap
c	26	46	3.2	402	9	US-09-864-761-3681	Sequence 3681, Ap
c	27	45	3.1	137	9	US-09-864-761-20449	Sequence 20449, A
c	28	45	3.1	221	9	US-09-893-238-7	Sequence 7, Appli
	29	33	2.3	5973	9	US-09-893-238-4	Sequence 4, Appli
c	30	30	2.1	17056	9	US-09-893-238-3	Sequence 3, Appli
	31	29	2.0	6733	16	US-10-197-824-1	Sequence 1, Appli
	32	28	2.0	594	15	US-10-027-632-264639	Sequence 264639,
	33	28	2.0	594	15	US-10-027-632-264640	Sequence 264640,
	34	26	1.8	629	15	US-10-027-632-270240	Sequence 270240,
	35	22	1.5	415	9	US-09-833-381-1916	Sequence 1916, Ap
c	36	22	1.5	5312	14	US-10-311-455-170	Sequence 170, App
	37	20	1.4	60	10	US-09-908-975-15421	Sequence 15421, A
	38	18	1.3	65	10	US-09-908-975-29990	Sequence 29990, A
	39	16	1.1	233	9	US-09-728-446-450	Sequence 450, App
c	40	14	1.0	2572	9	US-09-925-300-486	Sequence 486, App
c	41	14	1.0	3897	14	US-10-198-846-13651	Sequence 13651, A
c	42	14	1.0	3901	12	US-10-240-425-1425	Sequence 1425, Ap
c	43	14	1.0	3901	14	US-10-177-293-440	Sequence 440, App
c	44	14	1.0	3901	15	US-10-465-572-17	Sequence 17, Appl
c	45	14	1.0	3901	15	US-10-159-563-292	Sequence 292, App
c	46	14	1.0	3901	15	US-10-372-683-40	Sequence 40, Appl
c	47	14	1.0	225883	14	US-10-175-523-57	Sequence 57, Appl
	48	13	0.9	105	14	US-10-029-386-14368	Sequence 14368, A
c	49	13	0.9	520	15	US-10-027-632-181373	Sequence 181373,
	50	13	0.9	575	14	US-10-029-386-664	Sequence 664, App
	51	13	0.9	935	15	US-10-027-632-155757	Sequence 155757,
	52	13	0.9	2257	15	US-10-108-260A-876	Sequence 876, App
	53	13	0.9	2656	9	US-09-965-529-39	Sequence 39, Appl
	54	13	0.9	2656	10	US-09-969-680A-39	Sequence 39, Appl
	55	13	0.9	6691	14	US-10-301-822-3	Sequence 3, Appli
	56	13	0.9	8772	9	US-09-788-711A-3	Sequence 3, Appli
	57	13	0.9	8871	9	US-09-788-711A-1	Sequence 1, Appli
	58	13	0.9	9321	16	US-10-311-623-21	Sequence 21, Appl
	59	13	0.9	10531	14	US-10-225-567A-523	Sequence 523, App
	60	13	0.9	24370	15	US-10-292-798-931	Sequence 931, App
	61	13	0.9	68233	15	US-10-034-650-31	Sequence 31, Appl
c	62	12	0.8	287	9	US-09-864-761-24815	Sequence 24815, A
c	63	12	0.8	361	15	US-10-027-632-28416	Sequence 28416, A
c	64	12	0.8	394	14	US-10-198-846-13018	Sequence 13018, A
	65	12	0.8	411	13	US-10-040-739-599	Sequence 599, App
	66	12	0.8	436	9	US-09-864-761-2885	Sequence 2885, Ap
c	67	12	0.8	455	14	US-10-198-846-5901	Sequence 5901, Ap
c	68	12	0.8	465	10	US-09-930-213-151	Sequence 151, App
c	69	12	0.8	468	15	US-10-027-632-108988	Sequence 108988,
c	70	12	0.8	526	14	US-10-029-386-11510	Sequence 11510, A
c	71	12	0.8	575	9	US-09-864-761-8070	Sequence 8070, Ap
	72	12	0.8	641	14	US-10-101-464A-290	Sequence 290, App
	73	12	0.8	659	14	US-10-029-386-20268	Sequence 20268, A

c	74	12	0.8	834	12	US-10-424-599-17604	Sequence 17604, A
c	75	12	0.8	1256	12	US-10-425-114-35604	Sequence 35604, A
c	76	12	0.8	2023	15	US-10-264-049-540	Sequence 540, App
c	77	12	0.8	2322	15	US-10-027-632-110914	Sequence 110914,
c	78	12	0.8	2595	9	US-09-954-456-211	Sequence 211, App
c	79	12	0.8	2595	9	US-09-954-456-503	Sequence 503, App
	80	12	0.8	2751	9	US-09-728-422-2	Sequence 2, Appli
	81	12	0.8	3352	14	US-10-101-464A-868	Sequence 868, App
c	82	12	0.8	4020	14	US-10-198-846-12982	Sequence 12982, A
c	83	12	0.8	4320	15	US-10-440-464-139	Sequence 139, App
	84	11	0.8	125	9	US-09-864-761-28276	Sequence 28276, A
c	85	11	0.8	156	9	US-09-864-761-22124	Sequence 22124, A
	86	11	0.8	272	9	US-09-864-761-22608	Sequence 22608, A
	87	11	0.8	321	9	US-09-867-701-8623	Sequence 8623, Ap
	88	11	0.8	341	15	US-10-062-674-378	Sequence 378, App
	89	11	0.8	396	10	US-09-918-995-36657	Sequence 36657, A
c	90	11	0.8	434	9	US-09-783-590-9236	Sequence 9236, Ap
	91	11	0.8	454	9	US-09-864-761-5848	Sequence 5848, Ap
	92	11	0.8	454	10	US-09-918-995-24095	Sequence 24095, A
c	93	11	0.8	463	9	US-09-864-761-5350	Sequence 5350, Ap
	94	11	0.8	469	9	US-09-864-761-11698	Sequence 11698, A
c	95	11	0.8	492	14	US-10-029-386-25947	Sequence 25947, A
c	96	11	0.8	576	14	US-10-029-386-12247	Sequence 12247, A
	97	11	0.8	577	15	US-10-027-632-258512	Sequence 258512,
c	98	11	0.8	769	15	US-10-027-632-170724	Sequence 170724,
c	99	11	0.8	819	15	US-10-027-632-143212	Sequence 143212,
	100	11	0.8	876	9	US-09-740-668A-31	Sequence 31, Appl

ALIGNMENTS

RESULT 1
US-09-893-238-14
; Sequence 14, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 8589
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-893-238-14

Alignment Scores:

Pred. No.:	0	Length:	8589
Score:	1141.00	Matches:	1341
Percent Similarity:	99.70%	Conservative:	0
Best Local Similarity:	99.70%	Mismatches:	2
Query Match:	79.85%	Indels:	4
DB:	9	Gaps:	0

US-09-787-097-12 (1-1429) x US-09-893-238-14 (1-8589)

```
Qy        87 AlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGluCysAspArgProCysVal 106
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       256 GCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAAGGAATGTGACCGGCCGTGTGTC 315

Qy       107 AsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTrpValGly 126
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       316 AACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGC 375

Qy       127 GluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPheValThr 146
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       376 GAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACA 435

Qy       147 AspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeuIleGluGlyGlnPro 166
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       436 GATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCA 495

Qy       167 AsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLeu 186
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       496 AATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCATTTA 555

Qy       187 TyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaAlaPheSerGlyLeuIle 206
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       556 TATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATT 615

Qy       207 ValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGlyTyrAla 226
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       616 GTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCC 675

Qy       227 LeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThrTyrSer 246
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       676 TTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTACAGT 735

Qy       247 PheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSerSer 266
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       736 TTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGC 795

Qy       267 Glu-ThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIleProHi 286
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       796 GA-TACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCA 854

Qy       286 sCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspValArgGl 306
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       855 CTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGG 914

Qy       306 yCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAlaAsnGl 326
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       915 ATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCA 974

Qy       326 nSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHisLysAl 346
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       975 GTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGC 1034

Qy       346 aValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHisSerAspTy 366
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1035 TGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTA 1094

Qy       366 rAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSerVa 386
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1095 TAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGT 1154

Qy       386 lAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyrMe 406
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1155 GAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACAT 1214

Qy       406 tTyrGlyGlyLysIleAspPro-ThrGlyAsnValThrAsnGluLeuArgValPheHisI 426
             ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db      1215 GTATGGAGGAAAAATTGATTC-AACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACA 1273

Qy       426 leHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaValValG 446
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1274 TTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTG 1333

Qy       446 lyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValIlePheG 466
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1334 GGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTG 1393

Qy       466 lyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsnT 486
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1394 GTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACA 1453
```

```
Qy       486 hrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisSerServ 506
Db      1454 CATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTG 1513

Qy       506 alTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaA 526
Db      1514 TTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCCA 1573

Qy       526 snLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetTrpThrI 546
Db      1574 ATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCA 1633

Qy       546 leLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerGlyThrM 566
Db      1634 TTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACCA 1693

Qy       566 etLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysP 586
Db      1694 TGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATGCT 1753

Qy       586 heSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuProArgP 606
Db      1754 TCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGAC 1813

Qy       606 roAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrM 626
Db      1814 CTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACAACAGCACCA 1873

Qy       626 etTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSerG 646
Db      1874 TGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCGG 1933

Qy       646 luGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgC 666
Db      1934 AACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGT 1993

Qy       666 ysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnG 686
Db      1994 GTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAAG 2053

Qy       686 luGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArgCysAspG 706
Db      2054 AAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACC 2113

Qy       706 lnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAspH 726
Db      2114 AGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGCAATGACC 2173

Qy       726 isCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGluA 746
Db      2174 ATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGA 2233

Qy       746 snCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArgSerCysA 766
Db      2234 ATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTG 2293

Qy       766 laLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuProGluA 786
Db      2294 CCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAA 2353

Qy       786 snIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThrThrAlaL 806
Db      2354 ATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCA 2413

Qy       806 ysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSerL 826
Db      2414 AGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTC 2473

Qy       826 euThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGlnSerSerG 846
Db      2474 TTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTC 2533

Qy       846 lnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerT 866
Db      2534 AGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCT 2593

Qy       866 yrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerG 886
Db      2594 ACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTG 2653

Qy       886 luProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysA 906
Db      2654 AGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGG 2713

Qy       906 laAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSerA 926
Db      2714 CTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTG 2773

Qy       926 laLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGlyS 946
Db      2774 CTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGGCA 2833

Qy       946 erSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrValA 966
Db      2834 GCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGTGG 2893

Qy       966 laSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSerThrCysProProGluA 986
Db      2894 CCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGAAA 2953

Qy       986 snCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCysT 1006
Db      2954 ATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTA 3013

Qy      1006 hrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyProValL 1026
Db      3014 CTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGA 3073

Qy      1026 ysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerSerM 1046
Db      3074 AGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGCA 3133

Qy      1046 etCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnCysA 1066
Db      3134 TGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCA 3193

Qy      1066 snGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrThrG 1086
Db      3194 ACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCACAG 3253

Qy      1086 lyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyGlyLysC 1106
Db      3254 GCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAAT 3313

Qy      1106 ysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysP 1126
Db      3314 GTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGCT 3373

Qy      1126 heCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGluAsnArgT 1146
Db      3374 TCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGAT 3433

Qy      1146 yrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrGlnPheT 1166
Db      3434 ACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTTCA 3493

Qy      1166 hrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIleAsnPheValAlaThrP 1186
Db      3494 CCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTC 3553

Qy      1186 roAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeuA 1206
Db      3554 CTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCTCA 3613

Qy      1206 snIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProValV 1226
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3614 ACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTG 3673

Qy       1226 alSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPheAspPheA 1246
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3674 TTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTC 3733

Qy       1246 rgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpProIleLysI 1266
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3734 GCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAAAA 3793

Qy       1266 leGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePheValThrP 1286
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3794 TTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCTGGTACAGTTCTTCGTGACTT 3853

Qy       1286 hePheSerCysPheLeuSerLeuLeuLeuValAlaAlaValValTrpLysIleLysGlnS 1306
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3854 TCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGTGGTTTGGAAGATCAAACAAA 3913

Qy       1306 erCysTrpAlaSerArgArgArgGluGlnLeuLeuArgGluMetGlnGlnMetAlaSerA 1326
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3914 GTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGAGATGCAACAGATGGCCAGCC 3973

Qy       1326 rgProPheAlaSerValAsnValAlaLeuGluThrAspGluGluProProAspLeuIleG 1346
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3974 GTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTATTG 4033

Qy       1346 lyGlySerIleLysThrValProLysProIleAlaLeuGluProCysPheGlyAsnLysA 1366
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4034 GGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGAGCCGTGTTTTGGCAACAAAG 4093

Qy       1366 laAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyGlyIleProProProG 1386
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4094 CCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCTGGGTGGCATCCCTCCTCCTG 4153

Qy       1386 lyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetProIleV 1406
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4154 GGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGATAG 4213

Qy       1406 alTyrLysGluLysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnProG 1426
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4214 TGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGCCTG 4273

Qy       1426 lyThrCysIle 1429
                ||||||||||
Db       4274 GGACCTGCATC 4284

RESULT 2
US-09-893-238-16
; Sequence 16, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT:  Nagle, D.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
;  FILE REFERENCE: 7853-237
;  CURRENT APPLICATION NUMBER: US/09/893,238
;  CURRENT FILING DATE:  2001-06-27
;  PRIOR APPLICATION NUMBER: 09/245,041
;  PRIOR FILING DATE: 1999-02-05
;  PRIOR APPLICATION NUMBER: 60/093,630
;  PRIOR FILING DATE: 1998-07-21
;  PRIOR APPLICATION NUMBER: 60/104,978
;  PRIOR FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 16
;   LENGTH: 4072
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-893-238-16
```

```
Alignment Scores:
Pred. No.:                0                Length:          4072
Score:                    979.00           Matches:         1179
Percent Similarity:       99.66%           Conservative:    0
Best Lqcal Similarity:    99.66%           Mismatches:      2
Query Match:              68.51%           Indels:          4
DB:                       9                Gaps:            0

US-09-787-097-12 (1-1429) x US-09-893-238-16 (1-4072)

Qy         87 AlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGluCysAspArgProCysVal 106
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 GCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAAGGAATGTGACCGGCCGTGTGTC 315

Qy        107 AsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTrpValGly 126
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        316 AACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGC 375

Qy        127 GluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPheValThr 146
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        376 GAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACA 435

Qy        147 AspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeuIleGluGlyGlnPro 166
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        436 GATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCA 495

Qy        167 AsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLeu 186
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        496 AATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCATTTA 555

Qy        187 TyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaAlaPheSerGlyLeuIle 206
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 TATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATT 615

Qy        207 ValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGlyTyrAla 226
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 GTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCC 675

Qy        227 LeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThrTyrSer 246
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 TTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTACAGT 735

Qy        247 PheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSerSer 266
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 TTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGC 795

Qy        267 Glu-ThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIleProHi 286
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        796 GA-TACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCA 854

Qy        286 sCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspValArgGl 306
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 CTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGG 914

Qy        306 yCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAlaAsnGl 326
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 ATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCA 974

Qy        326 nSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHisLysAl 346
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 GTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGC 1034

Qy        346 aValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHisSerAspTy 366
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 TGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTA 1094

Qy        366 rAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSerVa 386
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 TAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGT 1154

Qy        386 lAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyrMe 406
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACAT 1214
```

```
Qy      406 tTyrGlyGlyLysIleAspPro-ThrGlyAsnValThrAsnGluLeuArgValPheHisI 426
Db     1215 GTATGGAGGAAAAATTGATTC-AACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACA 1273

Qy      426 leHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaValValG 446
Db     1274 TTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTG 1333

Qy      446 lyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValIlePheG 466
Db     1334 GGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTG 1393

Qy      466 lyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsnT 486
Db     1394 GTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACA 1453

Qy      486 hrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisSerServ 506
Db     1454 CATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTG 1513

Qy      506 alTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaA 526
Db     1514 TTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCCA 1573

Qy      526 snLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetTrpThrI 546
Db     1574 ATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCA 1633

Qy      546 leLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerGlyThrM 566
Db     1634 TTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACCA 1693

Qy      566 etLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysP 586
Db     1694 TGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATGCT 1753

Qy      586 heSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuProArgP 606
Db     1754 TCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGAC 1813

Qy      606 roAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrM 626
Db     1814 CTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACAACAGCACCA 1873

Qy      626 etTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSerG 646
Db     1874 TGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCGG 1933

Qy      646 luGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgC 666
Db     1934 AACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGT 1993

Qy      666 ysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnG 686
Db     1994 GTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAAG 2053

Qy      686 luGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArgCysAspG 706
Db     2054 AAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACC 2113

Qy      706 lnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAspH 726
Db     2114 AGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGCAATGACC 2173

Qy      726 isCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGluA 746
Db     2174 ATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGA 2233

Qy      746 snCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArgSerCysA 766
Db     2234 ATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTG 2293

Qy      766 laLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuProGluA 786
Db     2294 CCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAA 2353

Qy      786 snIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThrThrAlaL 806
Db     2354 ATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCA 2413

Qy      806 ysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSerL 826
Db     2414 AGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTC 2473

Qy      826 euThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGlnSerSerG 846
Db     2474 TTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTC 2533

Qy      846 lnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerT 866
Db     2534 AGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCT 2593

Qy      866 yrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerG 886
Db     2594 ACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTG 2653

Qy      886 luProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysA 906
Db     2654 AGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGG 2713

Qy      906 laAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSerA 926
Db     2714 CTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTG 2773

Qy      926 laLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGlyS 946
Db     2774 CTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGGCA 2833

Qy      946 erSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrValA 966
Db     2834 GCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGTGG 2893

Qy      966 laSerProProPheGlyGlnCysMetGluTrpTyrThrMetSerThrCysProProGluA 986
Db     2894 CCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGAAA 2953

Qy      986 snCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCysT 1006
Db     2954 ATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTA 3013

Qy     1006 hrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyProValL 1026
Db     3014 CTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGA 3073

Qy     1026 ysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerSerM 1046
Db     3074 AGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGCA 3133

Qy     1046 etCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnCysA 1066
Db     3134 TGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCA 3193

Qy     1066 snGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrThrG 1086
Db     3194 ACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCACAG 3253

Qy     1086 lyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyGlyLysC 1106
Db     3254 GCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAAT 3313

Qy     1106 ysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysP 1126
Db     3314 GTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGCT 3373

Qy     1126 heCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGluAsnArgT 1146
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3374 TCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGAT 3433

Qy       1146 yrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrGlnPheT 1166
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3434 ACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTTCA 3493

Qy       1166 hrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIleAsnPheValAlaThrP 1186
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3494 CCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTC 3553

Qy       1186 roAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeuA 1206
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3554 CTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCTCA 3613

Qy       1206 snIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProValV 1226
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3614 ACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTG 3673

Qy       1226 alSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPheAspPheA 1246
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3674 TTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTC 3733

Qy       1246 rgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpProIleLysI 1266
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3734 GCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAAAA 3793

Qy       1266 leGln 1267
                |||||
Db       3794 TTCAG 3798

RESULT 3
US-09-893-238-12
; Sequence 12, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT:  Nagle, D.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
;  FILE REFERENCE: 7853-237
;  CURRENT APPLICATION NUMBER: US/09/893,238
;  CURRENT FILING DATE:  2001-06-27
;  PRIOR APPLICATION NUMBER: 09/245,041
;  PRIOR FILING DATE: 1999-02-05
;  PRIOR APPLICATION NUMBER: 60/093,630
;  PRIOR FILING DATE: 1998-07-21
;  PRIOR APPLICATION NUMBER: 60/104,978
;  PRIOR FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 12
;   LENGTH: 6370
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-893-238-12

Alignment Scores:
Pred. No.:                0              Length:          6370
Score:                    503.00         Matches:         516
Percent Similarity:       99.81%         Conservative:    0
Best Local Similarity:    99.81%         Mismatches:      0
Query Match:              35.20%         Indels:          1
DB:                       9              Gaps:            0

US-09-787-097-12 (1-1429) x US-09-893-238-12 (1-6370)

Qy        914 AsnGlySerValCysGluArgProAlaAsnHisSerAlaLys-GlnCysArgThrProCy 933
                ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        506 AATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGGCAGTGCCGGACACCATG 565

Qy        933 sAlaLeuArgThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSe 953
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 TGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAG 625

Qy        953 rAsnMetLysGlnCysValAspSerAsnAlaTyrValAlaSerProPheProGlyGlnCy 973
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 CAACATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTG 685

Qy        973 sMetGluTrpTyrThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCy 993
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 TATGGAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTG 745

Qy        993 sSerHisCysLeuGluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLy 1013
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 TAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAA 805

Qy       1013 sGlyLysCysIleGluGlySerTyrLysGlyProValLysMetProSerGlnAlaProTh 1033
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 AGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTAC 865

Qy       1033 rGlyAsnPheTyrProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTy 1053
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 AGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATA 925

Qy       1053 rAsnTrpSerPheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAs 1073
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 CAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAA 985

Qy       1073 nGlnSerIleCysGluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIl 1093
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 TCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCAT 1045

Qy       1093 eSerGlyPheTyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGl 1113
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1046 ATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGG 1105

Qy       1113 yHisAlaSerLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLy 1133
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1106 GCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAA 1165

Qy       1133 sGlyAspGluCysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGl 1153
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1166 GGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGG 1225

Qy       1153 yThrCysTyrTyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAs 1173
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1226 AACATGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGA 1285

Qy       1173 pAspArgTyrTyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLe 1193
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1286 TGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTT 1345

Qy       1193 uAspMetPheIleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPh 1213
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1346 GGACATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTT 1405

Qy       1213 eSerAlaGlyThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGl 1233
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1406 CTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGA 1465

Qy       1233 uTyrLysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPh 1253
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1466 GTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTT 1525

Qy       1253 ePheValTyrValSerAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHi 1273
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1526 CTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGATTGCCTTCTCTCAGCA 1585

Qy       1273 sSerAsnPheMetAspLeuValGlnPhePheValThrPhePheSerCysPheLeuSerLe 1293
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1586 CAGCAATTTTATGGACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCTTT 1645

Qy       1293 uLeuLeuValAlaAlaValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgAr 1313
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1646 GCTCCTGGTGGCTGCTGTGGTTTGGAAGATCAAACAAAGTTGTTGGGCCTCCAGACGTAG 1705
```

```
Qy       1313 gGluGlnLeuLeuArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnVa 1333
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1706 AGAGCAACTTCTTCGAGAGATGCAACAGATGGCCAGCCGTCCCTTTGCCTCTGTAAATGT 1765

Qy       1333 lAlaLeuGluThrAspGluGluProProAspLeuIleGlyGlySerIleLysThrValPr 1353
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1766 CGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCC 1825

Qy       1353 oLysProIleAlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheVa 1373
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1826 CAAACCCATTGCACTGGAGCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGT 1885

Qy       1373 lArgLeuProArgGlyLeuGlyGlyIleProProProGlyGlnSerGlyLeuAlaValAl 1393
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1886 GAGGCTCCCTCGAGGCCTGGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGC 1945

Qy       1393 aSerAlaLeuValAspIleSerGlnGlnMetProIleValTyrLysGluLysSerGlyAl 1413
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1946 CAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGC 2005

Qy       1413 aValArgAsnArgLysGlnGlnProProAlaGlnProGlyThrCysIle 1429
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2006 CGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGCCTGGGACCTGCATC 2054
```

RESULT 4
US-09-893-238-18
; Sequence 18, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 2625
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-893-238-18

Alignment Scores:

Pred. No.:	0	Length:	2625
Score:	496.00	Matches:	696
Percent Similarity:	99.43%	Conservative:	0
Best Local Similarity:	99.43%	Mismatches:	2
Query Match:	34.71%	Indels:	4
DB:	9	Gaps:	0

US-09-787-097-12 (1-1429) x US-09-893-238-18 (1-2625)

```
Qy         87 AlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGluCysAspArgProCysVal 106
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 GCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAAGGAATGTGACCGGCCGTGTGTC 315

Qy        107 AsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTrpValGly 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        316 AACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGC 375

Qy        127 GluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPheValThr 146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        376 GAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACA 435

Qy        147 AspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeuIleGluGlyGlnPro 166
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        436 GATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCA 495

Qy        167 AsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLeu 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        496 AATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCATTTA 555

Qy        187 TyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaAlaPheSerGlyLeuIle 206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 TATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATT 615

Qy        207 ValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGlyTyrAla 226
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 GTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCC 675

Qy        227 LeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThrTyrSer 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 TTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTACAGT 735

Qy        247 PheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSerSer 266
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 TTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGC 795

Qy        267 Glu-ThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIleProHi 286
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        796 GA-TACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCA 854

Qy        286 sCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspValArgGl 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 CTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGG 914

Qy        306 yCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAlaAsnGl 326
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 ATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCA 974

Qy        326 nSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHisLysAl 346
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 GTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGC 1034

Qy        346 aValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHisSerAspTy 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 TGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTA 1094

Qy        366 rAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArgServa 386
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 TAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGT 1154

Qy        386 lAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyrMe 406
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACAT 1214

Qy        406 tTyrGlyGlyLysIleAspPro-ThrGlyAsnValThrAsnGluLeuArgValPheHisI 426
              |||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db       1215 GTATGGAGGAAAAATTGATTC-AACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACA 1273

Qy        426 leHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaValValG 446
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1274 TTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTG 1333

Qy        446 lyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValIlePheG 466
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1334 GGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTG 1393

Qy        466 lyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsnT 486
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1394 GTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACA 1453

Qy        486 hrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisSerServ 506
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1454 CATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTG 1513

Qy        506 alTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaA 526
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1514 TTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCCA 1573

Qy        526 snLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetTrpThrI 546
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1574 ATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCA 1633

Qy        546 leLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerGlyThrM 566
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1634 TTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACCA 1693

Qy        566 etLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysP 586
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1694 TGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATGCT 1753

Qy        586 heSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuProArgP 606
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1754 TCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGAC 1813

Qy        606 roAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrM 626
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1814 CTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACAACAGCACCA 1873

Qy        626 etTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSerG 646
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1874 TGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCGG 1933

Qy        646 luGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgC 666
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1934 AACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGT 1993

Qy        666 ysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnG 686
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1994 GTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAAG 2053

Qy        686 luGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArgCysAspG 706
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2054 AAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACC 2113

Qy        706 lnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAspH 726
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2114 AGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGCAATGACC 2173

Qy        726 isCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGluA 746
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2174 ATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGA 2233

Qy        746 snCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArgSerCysA 766
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2234 ATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTG 2293

Qy        766 laLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuPro 784
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2294 CCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCC 2349

RESULT 5
US-09-893-238-1
; Sequence 1, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT:  Nagle, D.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
;  FILE REFERENCE: 7853-237
;  CURRENT APPLICATION NUMBER: US/09/893,238
;  CURRENT FILING DATE:  2001-06-27
;  PRIOR APPLICATION NUMBER: 09/245,041
;  PRIOR FILING DATE: 1999-02-05
;  PRIOR APPLICATION NUMBER: 60/093,630
;  PRIOR FILING DATE: 1998-07-21
;  PRIOR APPLICATION NUMBER: 60/104,978
;  PRIOR FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 8827
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-893-238-1

Alignment Scores:
Pred. No.:                1.92e-193       Length:          8827
Score:                    206.00          Matches:         576
Percent Similarity:       96.00%          Conservative:    0
Best Local Similarity:    96.00%          Mismatches:      12
Query Match:              14.42%          Indels:          24
DB:                       9               Gaps:            0

US-09-787-097-12 (1-1429) x US-09-893-238-1 (1-8827)

Qy         842 MetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLys 861
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2790 ATGCAATCATCTCAAAGTATGTCCAAGCTCACTCTGACTCCATGGGTTGGTCTTCGGAAG 2849

Qy         862 IleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGln 881
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2850 ATCAATGTGTCTTACTGGTGCTGGGAGGATATGTCTCCATTCACAAATAGTTTGCTGCAG 2909

Qy         882 TrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThr 901
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2910 TGGATGCCATCTGAGCCCAGTGATGCTGGCTTCTGTGGGATCTTGTCAGAGCCTAGTACT 2969

Qy         902 ArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro 921
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2970 CGGGGATTAAAGGCTGCAACCTGCATCAACCCTCTCAATGGCAGCGTCTGTGAAAGGCCT 3029

Qy         922 AlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAsp 941
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3030 GCAAACCACAGTGCCAAGCAGTGCCGGACACCATGTGCCCTGCGGACAGCGTGTGGCGA- 3088

Qy         942 -CysThrSerGly-SerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspS 961
                 ||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3089 GTGCACTAGCAG-CAGCTCGGAGTGCATGTGGTGCAGTAACATGAAGCAGTGTGTGGACT 3147

Qy         961 erAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSerT 981
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3148 CCAATGCCTACGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCA 3207

Qy         981 hr-CysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPro 1000
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3208 G-CTGCCCACCTGAAAATTGCTCTGGCTACTGTACCTGCAGCCATTGCTTGGAGCAGCCA 3266

Qy        1001 GlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySer 1020
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3267 GGCTGTGGTTGGTGTACTGATCCTAGCAATACTGGGAAAGGAAAATGTATTGAGGGCAGC 3326

Qy        1021 TyrLysGlyProValLysMetProSerGlnAlaProThr--GlyAsnPhe-TyrProGln 1039
                 ||||||||||||||||||||||||||||||||||||||  ||||||  ||||||||
Db        3327 TATAAAGGACCTGTGAAGATGCCGTCACAGGCCTC-TG-CAGGAAATGT-GTATCCACAG 3383

Qy        1040 ProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHis 1059
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3384 CCCCTTCTGAACTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCAC 3443

Qy        1060 CysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLys 1079
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3444 TGTCCAGCTTGCCAGTGCAACGGACACAGCAAATGCATCAACCAGAGTATCTGTGAGAAG 3503

Qy        1080 CysGluAsn-LeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAs 1099
                 ||||||    ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3504 TGTGAGGA-CCTGACCACGGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTATGGTGA 3562

Qy        1099 pProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAs 1119
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3563 CCCGACTAATGGAGGCAAATGTCAGCCATGCAAGTGCAATGGGCACGCATCACTGTGCAA 3622
```

```
Qy       1119 nThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLe 1139
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3623 CACCAACACCGGCAAGTGCTTCTGTACCACCAAAGGTGTCAAGGGGGACGAGTGCCAGCT 3682

Qy       1139 uCysGluValGluAsnArgTyrGlnGlyAsnProLeuArg-GlyThrCysTyrTyrThrL 1159
             ||||||||||||||||||||||||||||||||||    |||||||||||||||||||||||
Db       3683 ATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAA-AGGAACATGCTACTATACCC 3741

Qy       1159 euLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrA 1179
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3742 TTCTCATTGACTATCAGTTCACCTTTAGCCTGTCCCAGGAAGACGACCGCTACTACACAG 3801

Qy       1179 laIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnA 1199
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3802 CCATCAACTTTGTGGCTACTCCTGATGAACAAAACAGGGATTTGGACATGTTCATCAATG 3861

Qy       1199 laSerLysAsnPheAsnLeuAsnIleThrTrpAlaAla-SerPheSer-AlaGlyThrGl 1218
             ||||||||||||||||||||||||||||||||||     ||||||    ||||||||||||
Db       3862 CCTCCAAAAACTTCAACCTCAACATCACCTGGGCCAC-CAGCTTCCC-AGCCGGAACCCA 3919

Qy       1218 nAla-GlyGluGluMet-ProValValSerLysThrAsnIleLysGluTyrLysAspSer 1237
             |    ||||||||    ||||||||||||||||||||||||||||||||||||||||||||
Db       3920 GAC-TGGAGAAGAGGT-GCCTGTTGTTTCAAAAACCAACATCAAGGAATACAAAGATAGC 3977

Qy       1238 PheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrVal 1257
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3978 TTCTCTAATGAGAAATTTGATTTTCGCAACCATCCAAACATCACTTTCTTTGTTTATGTC 4037

Qy       1258 SerAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMet 1277
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4038 AGTAATTTCACTTGGCCCATCAAAATTCAGATTGCCTTCTCCCAGCACAGCAACTTCATG 4097

Qy       1278 AspLeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuLeuValAla 1297
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4098 GACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCGCTGCTTCTGGTGGCT 4157

Qy       1298 AlaValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeuLeu 1317
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4158 GCAGTGGTCTGGAAGATCAAGCAGAGCTGTTGGGCATCCAGGCGGAGAGAGCAACTTCTT 4217

Qy       1318 ArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThr 1337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4218 CGGGAGATGCAACAGATGGCCAGCCGCCCCTTTGCTTCTGTAAACGTTGCCTTGGAAACA 4277

Qy       1338 AspGluGluProProAspLeuIleGlyGlySerIleLysThrValProLysProIleAla 1357
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4278 GATGAGGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCA 4337

Qy       1358 LeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProArg 1377
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4338 CTGGAGCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGA 4397

Qy       1378 GlyLeuGlyGlyIleProProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuVal 1397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4398 GGCCTGGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTG 4457

Qy       1398 AspIleSerGlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAsnArg 1417
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4458 GACATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGG 4517

Qy       1418 LysGlnGlnProProAlaGlnProGlyThrCysIle 1429
             |||||||||||||||||||||||||||||||||||||
Db       4518 AAGCAGCAGCCCCCTGCACAGCCTGGGACCTGCATC 4553

RESULT 6
US-09-893-238-8
; Sequence 8, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT:  Nagle, D.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
;  FILE REFERENCE: 7853-237
;  CURRENT APPLICATION NUMBER: US/09/893,238
;  CURRENT FILING DATE:  2001-06-27
;  PRIOR APPLICATION NUMBER: 09/245,041
;  PRIOR FILING DATE: 1999-02-05
;  PRIOR APPLICATION NUMBER: 60/093,630
;  PRIOR FILING DATE: 1998-07-21
;  PRIOR APPLICATION NUMBER: 60/104,978
;  PRIOR FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
;   LENGTH: 2419
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-893-238-8

Alignment Scores:
Pred. No.:                9.14e-83        Length:          2419
Score:                    94.00           Matches:         94
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              6.58%           Indels:          0
DB:                       9               Gaps:            0

US-09-787-097-12 (1-1429) x US-09-893-238-8 (1-2419)

Qy        513 AlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAsp 532
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1803 GCTCTGTACGTTCATGGTGGCTACAAGGCTTTCAGCGCCAACAAATACCGGCTTGCAGAT 1862

Qy        533 AspLeuTyrArgTyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPhe 552
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1863 GACCTCTACAGATACGATGTGGATACTCAGATGTGGACCATTCTTAAGGACAGCCGATTT 1922

Qy        553 PheArgTyrLeuHisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsn 572
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1923 TTCCGTTACTTGCATACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGAGGGAAC 1982

Qy        573 ThrHisAsnAspThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAla 592
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1983 ACACACAATGACACTTCCATGAGCCACGGTGCCAAATGCTTCTCCTCGGACTTCATGGCT 2042

Qy        593 TyrAspIleAlaCysAspArgTrpSerValLeuProArgPro 606
              ||||||||||||||||||||||||||||||||||||||||||
Db       2043 TATGACATTGCTTGTGACCGATGGTCAGTGCTTCCCAGACCT 2084

RESULT 7
US-10-198-846-5869/c
; Sequence 5869, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
;  APPLICANT: Lillie, James
;  APPLICANT:  Xu, Yongyao
;  APPLICANT:  Wang, Youzhen
;  APPLICANT:  Steinmann, Kathleen
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
;  TITLE OF INVENTION:  FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
;  TITLE OF INVENTION:  THERAPY OF BREAST CANCER
;  FILE REFERENCE: MRI-049
;  CURRENT APPLICATION NUMBER: US/10/198,846
;  CURRENT FILING DATE:  2002-07-18
;  PRIOR APPLICATION NUMBER: 60/306,220
;  PRIOR FILING DATE: 2001-07-18
;  NUMBER OF SEQ ID NOS: 14084
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5869
;   LENGTH: 625
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
```

```
;   LOCATION: 3, 4, 5, 6, 7, 8, 10, 12, 14, 15, 17, 26, 28, 29
;   OTHER INFORMATION: n = A,T,C or G
US-10-198-846-5869

Alignment Scores:
Pred. No.:              2.48e-81        Length:         625
Score:                  92.00           Matches:        92
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            6.44%           Indels:         0
DB:                     14              Gaps:           0

US-09-787-097-12 (1-1429) x US-10-198-846-5869 (1-625)

Qy        901 ThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArg 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 ACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGG 279

Qy        921 ProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGly 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 CCTGCAAACCACAGTGCTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGA 219

Qy        941 AspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAsp 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 GATTGCACCAGCGGCAGCTCTGAGTGTATGTGGTGCAGCAACATGAAGCAGTGTGTGGAC 159

Qy        961 SerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSer 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 TCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGC 99

Qy        981 ThrCysProProGluAsnCysSerGlyTyrCysThr 992
              ||||||||||||||||||||||||||||||||
Db         98 ACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACC 63

RESULT 8
US-09-736-457-966/c
; Sequence 966, Application US/09736457
; Patent No. US20020168637A1
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Wang, Aijun
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.478C15
;  CURRENT APPLICATION NUMBER: US/09/736,457
;  CURRENT FILING DATE:  2000-12-13
;  NUMBER OF SEQ ID NOS:  1864
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 966
;   LENGTH: 246
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-736-457-966

Alignment Scores:
Pred. No.:              8.74e-72        Length:         246
Score:                  82.00           Matches:        82
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            5.74%           Indels:         0
DB:                     9               Gaps:           0

US-09-787-097-12 (1-1429) x US-09-736-457-966 (1-246)

Qy        839 LeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGly 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy        859 LeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSer 878
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy        879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGlu 898
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67

Qy        899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy        919 GluArg 920
              ||||||
Db          6 GAAAGG 1

RESULT 9
US-09-902-941-966/c
; Sequence 966, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
;  APPLICANT: Henderson, Robert A.
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Watanabe, Yoshihiro
;  APPLICANT:  Johnson, Jeffrey C.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Marnerakis, Margarita
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  McNabb, Andria
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.478C17
;  CURRENT APPLICATION NUMBER: US/09/902,941
;  CURRENT FILING DATE:  2001-07-10
;  NUMBER OF SEQ ID NOS: 2002
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 966
;   LENGTH: 246
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-902-941-966

Alignment Scores:
Pred. No.:              8.74e-72        Length:         246
Score:                  82.00           Matches:        82
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            5.74%           Indels:         0
DB:                     9               Gaps:           0

US-09-787-097-12 (1-1429) x US-09-902-941-966 (1-246)

Qy        839 LeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGly 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy        859 LeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSer 878
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy        879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGlu 898
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67

Qy        899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7
```

```
Qy        919 GluArg 920
              ||||||
Db          6 GAAAGG 1

RESULT 10
US-09-849-626-966/c
; Sequence 966, Application US/09849626
; Publication No. US20020197669A1
; GENERAL INFORMATION:
;  APPLICANT: Bangur, Chaitanya
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Switzer, Anne
;  APPLICANT:  McNeill, Patricia
;  APPLICANT:  Clapper, Jonathan
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.478C16
;  CURRENT APPLICATION NUMBER: US/09/849,626
;  CURRENT FILING DATE:  2001-05-03
;  NUMBER OF SEQ ID NOS: 1926
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 966
;   LENGTH: 246
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-849-626-966

Alignment Scores:
Pred. No.:              8.74e-72        Length:         246
Score:                  82.00           Matches:        82
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            5.74%           Indels:         0
DB:                     9               Gaps:           0

US-09-787-097-12 (1-1429) x US-09-849-626-966 (1-246)

Qy        839 LeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGly 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy        859 LeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSer 878
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy        879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGlu 898
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67

Qy        899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy        919 GluArg 920
              ||||||
Db          6 GAAAGG 1

RESULT 11
US-10-017-754-966/c
; Sequence 966, Application US/10017754
; Publication No. US20030054363A1
; GENERAL INFORMATION:
;  APPLICANT: Henderson, Robert A.
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Watanabe, Yoshihiro
;  APPLICANT:  Johnson, Jeffrey C.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Marnerakis, Margarita
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  McNabb, Andria
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.478C18
;  CURRENT APPLICATION NUMBER: US/10/017,754
;  CURRENT FILING DATE:  2001-10-29
;  NUMBER OF SEQ ID NOS: 2004
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 966
;   LENGTH: 246
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-017-754-966

Alignment Scores:
Pred. No.:              8.74e-72        Length:         246
Score:                  82.00           Matches:        82
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            5.74%           Indels:         0
DB:                     14              Gaps:           0

US-09-787-097-12 (1-1429) x US-10-017-754-966 (1-246)

Qy        839 LeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGly 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy        859 LeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSer 878
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy        879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGlu 898
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67

Qy        899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy        919 GluArg 920
              ||||||
Db          6 GAAAGG 1

RESULT 12
US-10-113-872-966/c
; Sequence 966, Application US/10113872
; Publication No. US20030170255A1
; GENERAL INFORMATION:
;  APPLICANT: Watanabe, Yoshihiro
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.478C19
;  CURRENT APPLICATION NUMBER: US/10/113,872
;  CURRENT FILING DATE:  2002-03-28
;  NUMBER OF SEQ ID NOS: 2011
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 966
;   LENGTH: 246
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-113-872-966

Alignment Scores:
Pred. No.:              8.74e-72        Length:         246
```

Score:	82.00	Matches:	82
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.74%	Indels:	0
DB:	14	Gaps:	0

US-09-787-097-12 (1-1429) x US-10-113-872-966 (1-246)

```
Qy        839 LeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGly 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy        859 LeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSer 878
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy        879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGlu 898
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67

Qy        899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy        919 GluArg 920
              ||||||
Db          6 GAAAGG 1
```

RESULT 13
US-10-277-216-5
; Sequence 5, Application US/10277216
; Publication No. US20040002470A1
; GENERAL INFORMATION:
; APPLICANT: KEITH, TIM
; TITLE OF INVENTION: NOVEL HUMAN GENE RELATING TO RESPIRATORY DISEASES,
; TITLE OF INVENTION: OBESITY, AND INFLAMMATORY BOWEL DISEASE
; FILE REFERENCE: 2976-4051
; CURRENT APPLICATION NUMBER: US/10/277,216
; CURRENT FILING DATE: 2002-10-17
; PRIOR APPLICATION NUMBER: 10/126,022
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 09/834,597
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 09/548,797
; PRIOR FILING DATE: 2000-04-13
; NUMBER OF SEQ ID NOS: 420
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 207433
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-277-216-5

Alignment Scores:

Pred. No.:	4.58e-64	Length:	207433
Score:	77.00	Matches:	77
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.39%	Indels:	0
DB:	15	Gaps:	0

US-09-787-097-12 (1-1429) x US-10-277-216-5 (1-207433)

```
Qy         985 GluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrp 1004
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        9815 GAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGG 9874

Qy        1005 CysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyPro 1024
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        9875 TGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCA 9934

Qy        1025 ValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSer 1044
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        9935 GTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCC 9994

Qy        1045 SerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysPro 1061
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        9995 AGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCA 10045
```

RESULT 14
US-10-126-022-5
; Sequence 5, Application US/10126022
; Publication No. US20040023215A1
; GENERAL INFORMATION:
; APPLICANT: KEITH, TIM
; TITLE OF INVENTION: NOVEL HUMAN GENE RELATING TO RESPIRATORY DISEASES,
; TITLE OF INVENTION: OBESITY, AND INFLAMMATORY BOWEL DISEASE
; FILE REFERENCE: 2976-4039US2
; CURRENT APPLICATION NUMBER: US/10/126,022
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 09/834,597
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 09/548,797
; PRIOR FILING DATE: 2000-04-13
; NUMBER OF SEQ ID NOS: 420
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 207433
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-126-022-5

Alignment Scores:

Pred. No.:	4.58e-64	Length:	207433
Score:	77.00	Matches:	77
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.39%	Indels:	0
DB:	16	Gaps:	0

US-09-787-097-12 (1-1429) x US-10-126-022-5 (1-207433)

```
Qy         985 GluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrp 1004
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        9815 GAAAATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGG 9874

Qy        1005 CysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyPro 1024
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        9875 TGTACTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCA 9934

Qy        1025 ValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSer 1044
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        9935 GTGAAGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCC 9994

Qy        1045 SerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysPro 1061
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        9995 AGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCA 10045
```

RESULT 15
US-09-893-238-10
; Sequence 10, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20

```
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10
;   LENGTH: 1051
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-893-238-10

Alignment Scores:
Pred. No.:                3.11e-65       Length:          1051
Score:                    76.00          Matches:         111
Percent Similarity:       98.23%         Conservative:    0
Best Local Similarity:    98.23%         Mismatches:      1
Query Match:              5.32%          Indels:          2
DB:                       9              Gaps:            0

US-09-787-097-12 (1-1429) x US-09-893-238-10 (1-1051)

Qy          87 AlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGluCysAspArgProCysVal 106
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         463 GCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAAGGAATGTGACCGGCCGTGTGTC 522

Qy         107 AsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCys-ProAlaGlyTrpValGl 126
               |||||||||||||||||||||||||||||||||||||||||| |||   ||||||||||
Db         523 AACGGCGGCCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCACG-GGCTGGGTGGG 581

Qy         126 yGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPheValTh 146
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         582 CGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGCTCTTCTGGATTTGTAAC 641

Qy         146 rAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeuIleGluGlyGlnPr 166
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         642 AGATGGACCTGGGAATTATAAATATAAGACGAAGTGCACATGGCTCATTGAAGGACAGCC 701

Qy         166 oAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLe 186
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         702 AAATAGAATAATGAGACTTCGCTTCAACCATTTTGCTACAGAATGTAGCTGGGACCATTT 761

Qy         186 uTyrValTyrAspGlyAspSerIleTyrAlaProLeu 198
               |||||||||||||||||||||||||||||||||||
Db         762 ATATGTTTATGATGGGGACTCAATCTACGCACCTCTG 798

RESULT 16
US-09-893-238-5
; Sequence 5, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT:  Nagle, D.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
;  FILE REFERENCE: 7853-237
;  CURRENT APPLICATION NUMBER: US/09/893,238
;  CURRENT FILING DATE:  2001-06-27
;  PRIOR APPLICATION NUMBER: 09/245,041
;  PRIOR FILING DATE: 1999-02-05
;  PRIOR APPLICATION NUMBER: 60/093,630
;  PRIOR FILING DATE: 1998-07-21
;  PRIOR APPLICATION NUMBER: 60/104,978
;  PRIOR FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5
;   LENGTH: 90050
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-893-238-5

Alignment Scores:
Pred. No.:                1.96e-60       Length:          90050
Score:                    73.00          Matches:         73
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              5.11%          Indels:          0
DB:                       9              Gaps:            0

US-09-787-097-12 (1-1429) x US-09-893-238-5 (1-90050)

Qy          849 SerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        28996 TCCAAGCTCACTCTGACTCCATGGGTTGGTCTTCGGAAGATCAATGTGTCTTACTGGTGC 29055

Qy          869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        29056 TGGGAGGATATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGT 29115

Qy          889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        29116 GATGCTGGCTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGATTAAAGGCTGCAACC 29175

Qy          909 CysIleAsnProLeuAsnGlySerValCysGluArgPro 921
                |||||||||||||||||||||||||||||||||||||
Db        29176 TGCATCAACCCTCTCAATGGCAGCGTCTGTGAAAGGCCT 29214

RESULT 17
US-09-864-761-15900
; Sequence 15900, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
```

```
; SEQ ID NO 15900
;   LENGTH: 531
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL132773.14
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
US-09-864-761-15900

Alignment Scores:
Pred. No.:              1.46e-58        Length:         531
Score:                  69.00           Matches:        69
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            4.83%           Indels:         0
DB:                     9               Gaps:           0

US-09-787-097-12 (1-1429) x US-09-864-761-15900 (1-531)

Qy        414 ThrGlyAsnValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeu 433
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 ACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTG 245

Qy        434 LeuThrProLysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThr 453
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 TTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACA 305

Qy        454 LeuLysAsnGlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyr 473
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 CTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATAT 365

Qy        474 IleSerAsnValGlnGluTyrAspLeu 482
              |||||||||||||||||||||||||||
Db        366 ATAAGCAATGTGCAGGAATATGATTTG 392

RESULT 18
US-09-864-761-32404
; Sequence 32404, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 32404
;   LENGTH: 193
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL132773.14
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
;   OTHER INFORMATION: EST_HUMAN HIT: BF313142.1, EVALUE 1.00e-104
;   OTHER INFORMATION: NT HIT: AF218896.1, EVALUE 1.00e-105
;   OTHER INFORMATION: SWISSPROT HIT: Q19981, EVALUE 1.00e-05
US-09-864-761-32404

Alignment Scores:
Pred. No.:              5.19e-54        Length:         193
Score:                  64.00           Matches:        64
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            4.48%           Indels:         0
DB:                     9               Gaps:           0

US-09-787-097-12 (1-1429) x US-09-864-761-32404 (1-193)

Qy        416 AsnValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThr 435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 AATGTGACCAATGAGTTGAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACC 61

Qy        436 ProLysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLys 455
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CCTAAGGCAAAGGAGCAGTATGCAGTGGTTGGGCACTCTGCACACATTGTTACACTGAAG 121

Qy        456 AsnGlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSer 475
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 AATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGC 181

Qy        476 AsnValGlnGlu 479
              ||||||||||||
Db        182 AATGTGCAGGAA 193

RESULT 19
US-09-864-761-21460/c
; Sequence 21460, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
```

```
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 21460
;   LENGTH: 182
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL132773.1
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.96
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL =   1
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.6
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.6
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL =   1
;   OTHER INFORMATION: NT HIT: AF218902.1, EVALUE 2.00e-98
;   OTHER INFORMATION: EST_HUMAN HIT: AI372810.1, EVALUE 2.00e-98
;   OTHER INFORMATION: SWISSPROT HIT: P22897, EVALUE 8.00e-03
US-09-864-761-21460

Alignment Scores:
Pred. No.:                4.6e-50        Length:          182
Score:                    60.00          Matches:         60
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              4.20%          Indels:          0
DB:                       9              Gaps:            0

US-09-787-097-12 (1-1429) x US-09-864-761-21460 (1-182)

Qy        849 SerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCTACTGGTGC 123

Qy        869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 TGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAGT 63

Qy        889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACC 3

RESULT 20
US-09-864-761-4720/c
; Sequence 4720, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 4720
;   LENGTH: 375
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL132773.1
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.96
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL =   1
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.6
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.6
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL =   1
US-09-864-761-4720
```

```
Alignment Scores:
Pred. No.:                8.65e-47        Length:         375
Score:                    57.00           Matches:        57
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              3.99%           Indels:         0
DB:                       9               Gaps:           0

US-09-787-097-12 (1-1429) x US-09-864-761-4720 (1-375)

Qy         865 SerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetPro 884
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         374 TCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCG 315

Qy         885 SerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeu 904
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         314 TCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTG 255

Qy         905 LysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro 921
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         254 AAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCT 204

RESULT 21
US-09-893-238-6/c
; Sequence 6, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT:  Nagle, D.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
;  FILE REFERENCE: 7853-237
;  CURRENT APPLICATION NUMBER: US/09/893,238
;  CURRENT FILING DATE:  2001-06-27
;  PRIOR APPLICATION NUMBER: 09/245,041
;  PRIOR FILING DATE: 1999-02-05
;  PRIOR APPLICATION NUMBER: 60/093,630
;  PRIOR FILING DATE: 1998-07-21
;  PRIOR APPLICATION NUMBER: 60/104,978
;  PRIOR FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 6
;   LENGTH: 226
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-893-238-6

Alignment Scores:
Pred. No.:                4.87e-41        Length:         226
Score:                    51.00           Matches:        51
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              3.57%           Indels:         0
DB:                       9               Gaps:           0

US-09-787-097-12 (1-1429) x US-09-893-238-6 (1-226)

Qy         871 AspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSerAspAla 890
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         225 GATATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGTGATGCT 166

Qy         891 GlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThrCysIle 910
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         165 GGCTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGATTAAAGGCTGCAACCTGCATC 106

Qy         911 AsnProLeuAsnGlySerValCysGluArgPro 921
               |||||||||||||||||||||||||||||||
Db         105 AACCCTCTCAATGGCAGCGTCTGTGAAAGGCCT 73

RESULT 22
US-09-796-692-7610
; Sequence 7610, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Mannion, Jane
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
;  TITLE OF INVENTION:  HEMATOLOGICAL MALIGNANCIES
;  FILE REFERENCE: 2077.001200
;  CURRENT APPLICATION NUMBER: US/09/796,692
;  CURRENT FILING DATE:  2001-03-01
;  PRIOR APPLICATION NUMBER: 60/186,126
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: 60/190,479
;  PRIOR FILING DATE: 2000-03-17
;  PRIOR APPLICATION NUMBER: 60/200,545
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 60/200,303
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: 60/200,779
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: 60/200,999
;  PRIOR FILING DATE: 2000-05-01
;  PRIOR APPLICATION NUMBER: 60/202,084
;  PRIOR FILING DATE: 2000-05-04
;  PRIOR APPLICATION NUMBER: 60/206,201
;  PRIOR FILING DATE: 2000-05-22
;  PRIOR APPLICATION NUMBER: 60/218,950
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 60/222,903
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: 60/223,416
;  PRIOR FILING DATE: 2000-08-04
;  PRIOR APPLICATION NUMBER: 60/223,378
;  PRIOR FILING DATE: 2000-08-07
;  NUMBER OF SEQ ID NOS: 9597
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7610
;   LENGTH: 212
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (140)
;   OTHER INFORMATION: n=A,T,C or G
;   NAME/KEY: unsure
;   LOCATION: (142)
;   OTHER INFORMATION: n=A,T,C or G
;   NAME/KEY: unsure
;   LOCATION: (162)
;   OTHER INFORMATION: n=A,T,C or G
;   NAME/KEY: unsure
;   LOCATION: (196)
;   OTHER INFORMATION: n=A,T,C or G
;   NAME/KEY: unsure
;   LOCATION: (204)
;   OTHER INFORMATION: n=A,T,C or G
;   NAME/KEY: unsure
;   LOCATION: (205)
;   OTHER INFORMATION: n=A,T,C or G
US-09-796-692-7610

Alignment Scores:
Pred. No.:                4.23e-36        Length:         212
Score:                    46.00           Matches:        46
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              3.22%           Indels:         0
DB:                       9               Gaps:           0

US-09-787-097-12 (1-1429) x US-09-796-692-7610 (1-212)

Qy         972 GlnCysMetGluTrpTyrThrMetSerThrCysProProGluAsnCysSerGlyTyrCys 991
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db          2 CAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGT 61

Qy        992 ThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCysThrAspProSerAsnThr 1011
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 ACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACT 121

Qy       1012 GlyLysGlyLysCysIle 1017
              |||||||||||||||||
Db        122 GGCAAAGGGAAATGCATA 139

RESULT 23
US-10-040-862-7610
; Sequence 7610, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
;  TITLE OF INVENTION:  Hematological Malignancies
;  FILE REFERENCE: 014058-013520US
;  CURRENT APPLICATION NUMBER: US/10/040,862
;  CURRENT FILING DATE:  2001-11-06
;  PRIOR APPLICATION NUMBER: US 60/186,126
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: US 60/190,479
;  PRIOR FILING DATE: 2000-03-17
;  PRIOR APPLICATION NUMBER: US 60/200,545
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: US 60/200,303
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: US 60/200,779
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: US 60/200,999
;  PRIOR FILING DATE: 2000-05-01
;  PRIOR APPLICATION NUMBER: US 60/202,084
;  PRIOR FILING DATE: 2000-05-04
;  PRIOR APPLICATION NUMBER: US 60/206,201
;  PRIOR FILING DATE: 2000-05-22
;  PRIOR APPLICATION NUMBER: US 60/218,950
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: US 60/222,903
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: US 60/223,416
;  PRIOR FILING DATE: 2000-08-04
;  PRIOR APPLICATION NUMBER: US 60/223,378
;  PRIOR FILING DATE: 2000-08-07
;  PRIOR APPLICATION NUMBER: US 09/796,692
;  PRIOR FILING DATE: 2001-03-01
;  NUMBER OF SEQ ID NOS: 10467
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7610
;   LENGTH: 212
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (140)
;   OTHER INFORMATION: n=A,T,C or G
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (142)
;   OTHER INFORMATION: n=A,T,C or G
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (162)
;   OTHER INFORMATION: n=A,T,C or G
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (196)
;   OTHER INFORMATION: n=A,T,C or G
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (204)
;   OTHER INFORMATION: n=A,T,C or G
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (205)
;   OTHER INFORMATION: n=A,T,C or G
US-10-040-862-7610

Alignment Scores:
Pred. No.:              4.23e-36        Length:         212
Score:                  46.00           Matches:        46
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            3.22%           Indels:         0
DB:                     14              Gaps:           0

US-09-787-097-12 (1-1429) x US-10-040-862-7610 (1-212)

Qy        972 GlnCysMetGluTrpTyrThrMetSerThrCysProProGluAsnCysSerGlyTyrCys 991
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 CAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGT 61

Qy        992 ThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCysThrAspProSerAsnThr 1011
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 ACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACT 121

Qy       1012 GlyLysGlyLysCysIle 1017
              |||||||||||||||||
Db        122 GGCAAAGGGAAATGCATA 139

RESULT 24
US-10-057-475B-7610
; Sequence 7610, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Clapper, Jonathan David
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Ordonez, Nadia
;  APPLICANT:  Carter, Lauren
;  APPLICANT:  McNeill, Patricia Dianne
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
;  TITLE OF INVENTION:  Hematological Malignancies
;  FILE REFERENCE: 014058-014402US
;  CURRENT APPLICATION NUMBER: US/10/057,475B
;  CURRENT FILING DATE:  2002-01-22
;  PRIOR APPLICATION NUMBER: US 60/186,126
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: US 60/190,479
;  PRIOR FILING DATE: 2000-03-17
;  PRIOR APPLICATION NUMBER: US 60/200,545
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: US 60/200,303
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: US 60/200,779
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: US 60/200,999
;  PRIOR FILING DATE: 2000-05-01
;  PRIOR APPLICATION NUMBER: US 60/202,084
;  PRIOR FILING DATE: 2000-05-04
;  PRIOR APPLICATION NUMBER: US 60/206,201
;  PRIOR FILING DATE: 2000-05-22
;  PRIOR APPLICATION NUMBER: US 60/218,950
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: US 60/222,903
;  PRIOR FILING DATE: 2000-08-03
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 10979
```

```
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7610
;   LENGTH: 212
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(212)
;   OTHER INFORMATION: n = g, a, c or t
US-10-057-475B-7610

Alignment Scores:
Pred. No.:                4.23e-36        Length:         212
Score:                    46.00           Matches:        46
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              3.22%           Indels:         0
DB:                       15              Gaps:           0

US-09-787-097-12 (1-1429) x US-10-057-475B-7610 (1-212)

Qy         972 GlnCysMetGluTrpTyrThrMetSerThrCysProProGluAsnCysSerGlyTyrCys 991
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           2 CAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGT 61

Qy         992 ThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCysThrAspProSerAsnThr 1011
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          62 ACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACT 121

Qy        1012 GlyLysGlyLysCysIle 1017
               ||||||||||||||||||
Db         122 GGCAAAGGGAAATGCATA 139

RESULT 25
US-10-154-884B-7610
; Sequence 7610, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
;  TITLE OF INVENTION:  Hematological Malignancies
;  FILE REFERENCE: 014058-013521US
;  CURRENT APPLICATION NUMBER: US/10/154,884B
;  CURRENT FILING DATE:  2002-05-23
;  PRIOR APPLICATION NUMBER: US 60/186,126
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: US 60/190,479
;  PRIOR FILING DATE: 2000-03-17
;  PRIOR APPLICATION NUMBER: US 60/200,545
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: US 60/200,303
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: US 60/200,779
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: US 60/200,999
;  PRIOR FILING DATE: 2000-05-01
;  PRIOR APPLICATION NUMBER: US 60/202,084
;  PRIOR FILING DATE: 2000-05-04
;  PRIOR APPLICATION NUMBER: US 60/206,201
;  PRIOR FILING DATE: 2000-05-22
;  PRIOR APPLICATION NUMBER: US 60/218,950
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: US 60/222,903
;  PRIOR FILING DATE: 2000-08-03
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 11290
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7610
;   LENGTH: 212
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(212)
;   OTHER INFORMATION: n = g, a, c or t
US-10-154-884B-7610

Alignment Scores:
Pred. No.:                4.23e-36        Length:         212
Score:                    46.00           Matches:        46
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              3.22%           Indels:         0
DB:                       15              Gaps:           0

US-09-787-097-12 (1-1429) x US-10-154-884B-7610 (1-212)

Qy         972 GlnCysMetGluTrpTyrThrMetSerThrCysProProGluAsnCysSerGlyTyrCys 991
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           2 CAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGT 61

Qy         992 ThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCysThrAspProSerAsnThr 1011
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          62 ACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACT 121

Qy        1012 GlyLysGlyLysCysIle 1017
               ||||||||||||||||||
Db         122 GGCAAAGGGAAATGCATA 139

RESULT 26
US-09-864-761-3681/c
; Sequence 3681, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
```

```
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 3681
;   LENGTH: 402
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL132773.1
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL =   2
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.5
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.1
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL =   2
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.7
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.1
US-09-864-761-3681

Alignment Scores:
Pred. No.:               7.72e-36       Length:         402
Score:                   46.00          Matches:        46
Percent Similarity:      100.00%        Conservative:   0
Best Local Similarity:   100.00%        Mismatches:     0
Query Match:             3.22%          Indels:         0
DB:                      9              Gaps:           0

US-09-787-097-12 (1-1429) x US-09-864-761-3681 (1-402)

Qy       1062 AlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGlu 1081
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 GCTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAG 330

Qy       1082 AsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThr 1101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 AACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACC 270

Qy       1102 AsnGlyGlyLysCysGln 1107
              |||||||||||||||||
Db        269 AATGGAGGGAAATGTCAG 252

RESULT 27
US-09-864-761-20449/c
; Sequence 20449, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 20449
;   LENGTH: 137
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL132773.1
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL =   2
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.5
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.1
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL =   2
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.7
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.1
;   OTHER INFORMATION: EST_HUMAN HIT: BF672370.1, EVALUE 2.00e-71
;   OTHER INFORMATION: NT HIT: AB011120.1, EVALUE 1.00e-71
;   OTHER INFORMATION: SWISSPROT HIT: Q19981, EVALUE 3.00e-11
US-09-864-761-20449

Alignment Scores:
Pred. No.:               2.76e-35       Length:         137
Score:                   45.00          Matches:        45
Percent Similarity:      100.00%        Conservative:   0
Best Local Similarity:   100.00%        Mismatches:     0
Query Match:             3.15%          Indels:         0
DB:                      9              Gaps:           0

US-09-787-097-12 (1-1429) x US-09-864-761-20449 (1-137)

Qy       1063 CysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsn 1082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 TGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAAC 76

Qy       1083 LeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsn 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAAT 16

Qy       1103 GlyGlyLysCysGln 1107
              |||||||||||||||
Db         15 GGAGGGAAATGTCAG 1

RESULT 28
US-09-893-238-7/c
```

```
; Sequence 7, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT:  Nagle, D.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
;  FILE REFERENCE: 7853-237
;  CURRENT APPLICATION NUMBER: US/09/893,238
;  CURRENT FILING DATE:  2001-06-27
;  PRIOR APPLICATION NUMBER: 09/245,041
;  PRIOR FILING DATE: 1999-02-05
;  PRIOR APPLICATION NUMBER: 60/093,630
;  PRIOR FILING DATE: 1998-07-21
;  PRIOR APPLICATION NUMBER: 60/104,978
;  PRIOR FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7
;   LENGTH: 221
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-893-238-7

Alignment Scores:
Pred. No.:              4.32e-35        Length:         221
Score:                  45.00           Matches:        45
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            3.15%           Indels:         0
DB:                     9               Gaps:           0

US-09-787-097-12 (1-1429) x US-09-893-238-7 (1-221)

Qy        871 AspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSerAspAla 890
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 GATATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGTGATGCT 161

Qy        891 GlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThrCysIle 910
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 GGCTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGATTAAAGGCTGCAACCTGCATC 101

Qy        911 AsnProLeuAsnGly 915
              |||||||||||||||
Db        100 AACCCTCTCAATGGT 86

RESULT 29
US-09-893-238-4
; Sequence 4, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT:  Nagle, D.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
;  FILE REFERENCE: 7853-237
;  CURRENT APPLICATION NUMBER: US/09/893,238
;  CURRENT FILING DATE:  2001-06-27
;  PRIOR APPLICATION NUMBER: 09/245,041
;  PRIOR FILING DATE: 1999-02-05
;  PRIOR APPLICATION NUMBER: 60/093,630
;  PRIOR FILING DATE: 1998-07-21
;  PRIOR APPLICATION NUMBER: 60/104,978
;  PRIOR FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
;   LENGTH: 5973
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-893-238-4

Alignment Scores:
Pred. No.:              7.92e-22        Length:         5973
Score:                  33.00           Matches:        33
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            2.31%           Indels:         0
DB:                     9               Gaps:           0

US-09-787-097-12 (1-1429) x US-09-893-238-4 (1-5973)

Qy        166 ProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHis 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4328 CCAAATAGAATAATGAGACTTCGCTTCAACCATTTTGCTACAGAATGTAGCTGGGACCAT 4387

Qy        186 LeuTyrValTyrAspGlyAspSerIleTyrAlaProLeu 198
              |||||||||||||||||||||||||||||||||||||||
Db       4388 TTATATGTTTATGATGGGGACTCAATCTACGCACCTCTG 4426

RESULT 30
US-09-893-238-3/c
; Sequence 3, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT:  Nagle, D.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
;  FILE REFERENCE: 7853-237
;  CURRENT APPLICATION NUMBER: US/09/893,238
;  CURRENT FILING DATE:  2001-06-27
;  PRIOR APPLICATION NUMBER: 09/245,041
;  PRIOR FILING DATE: 1999-02-05
;  PRIOR APPLICATION NUMBER: 60/093,630
;  PRIOR FILING DATE: 1998-07-21
;  PRIOR APPLICATION NUMBER: 60/104,978
;  PRIOR FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 17056
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-893-238-3

Alignment Scores:
Pred. No.:              2.03e-18        Length:         17056
Score:                  30.00           Matches:        30
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            2.10%           Indels:         0
DB:                     9               Gaps:           0

US-09-787-097-12 (1-1429) x US-09-893-238-3 (1-17056)

Qy        337 LeuLysLeuProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValVal 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2473 TTAAAGCTTCCCAGAGCCTCTCATAAAGCTGTGGTCAATGGAAATATAATGTGGGTTGTT 2414

Qy        357 GlyGlyTyrMetPheAsnHisSerAspTyr 366
              |||||||||||||||||||||||||||||||
Db       2413 GGCGGATATATGTTCAACCATTCAGATTAC 2384

RESULT 31
US-10-197-824-1
; Sequence 1, Application US/10197824
; Publication No. US20040023219A1
; GENERAL INFORMATION:
;  APPLICANT: ORIGENE TECHNOLOGIES INC
;  TITLE OF INVENTION: NOVEL PROSTATE CANCER GENES
;  FILE REFERENCE: 9U 106 R1
;  CURRENT APPLICATION NUMBER: US/10/197,824
;  CURRENT FILING DATE:  2002-07-19
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: PatentIn version 3.1
```

; SEQ ID NO 1
; LENGTH: 6733
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (575)..(4711)
; OTHER INFORMATION:
US-10-197-824-1

Alignment Scores:

Pred. No.:	8.31e-18	Length:	6733
Score:	29.00	Matches:	29
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.03%	Indels:	0
DB:	16	Gaps:	0

US-09-787-097-12 (1-1429) x US-10-197-824-1 (1-6733)

```
Qy       1277 MetAspLeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuLeuVal 1296
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4256 ATGGACCTTGTGCAGTTTTTTGTCACCTTCTTCAGTTGTTTCCTATCCTTATTGCTGGTG 4315

Qy       1297 AlaAlaValValTrpLysIleLysGln 1305
              |||||||||||||||||||||||||||
Db       4316 GCTGCTGTGGTATGGAAGATCAAACAA 4342
```

RESULT 32
US-10-027-632-264639
; Sequence 264639, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 264639
; LENGTH: 594
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-264639

Alignment Scores:

Pred. No.:	8.31e-18	Length:	594
Score:	28.00	Matches:	28
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.96%	Indels:	0
DB:	15	Gaps:	0

US-09-787-097-12 (1-1429) x US-10-027-632-264639 (1-594)

```
Qy        596 AlaCysAspArgTrpSerValLeuProArgProAspLeuHisHisAspValAsnArgPhe 615
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 GCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTT 203

Qy        616 GlyHisSerAlaValLeuHisAsn 623
              |||||||||||||||||||||||
Db        204 GGCCATTCAGCAGTCTTACACAAC 227
```

RESULT 33
US-10-027-632-264640
; Sequence 264640, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 264640
; LENGTH: 594
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-264640

Alignment Scores:

Pred. No.:	8.31e-18	Length:	594
Score:	28.00	Matches:	28
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.96%	Indels:	0
DB:	15	Gaps:	0

US-09-787-097-12 (1-1429) x US-10-027-632-264640 (1-594)

```
Qy        596 AlaCysAspArgTrpSerValLeuProArgProAspLeuHisHisAspValAsnArgPhe 615
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 GCCTGTGACCGCTGGTCAGTGCTTCCCAGACCTGATCTCCACCATGATGTCAACAGATTT 203

Qy        616 GlyHisSerAlaValLeuHisAsn 623
              |||||||||||||||||||||||
Db        204 GGCCATTCAGCAGTCTTACACAAC 227
```

RESULT 34
US-10-027-632-270240
; Sequence 270240, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483

; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 270240
; LENGTH: 629
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-270240

Alignment Scores:

Pred. No.:	8.49e-16	Length:	629
Score:	26.00	Matches:	26
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.82%	Indels:	0
DB:	15	Gaps:	0

US-09-787-097-12 (1-1429) x US-10-027-632-270240 (1-629)

```
Qy       1189 GlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeuAsnIleThr 1208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 CAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACC 406

Qy       1209 TrpAlaAlaSerPheSer 1214
              ||||||||||||||||||
Db        407 TGGGCTGCCAGTTTCTCA 424
```

RESULT 35
US-09-833-381-1916
; Sequence 1916, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1916
; LENGTH: 415
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(415)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1916

Alignment Scores:

Pred. No.:	5.38e-12	Length:	415
Score:	22.00	Matches:	22
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.54%	Indels:	0
DB:	9	Gaps:	0

US-09-787-097-12 (1-1429) x US-09-833-381-1916 (1-415)

```
Qy        222 ThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPhe 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 ACATCTGGCTATGCACTGTTACATTTTTTTAGTGATGCTGCGTATAATCTAACTGGTTTC 306

Qy        242 AsnIle 243
              ||||||
Db        307 AACATT 312
```

RESULT 36
US-10-311-455-170/c
; Sequence 170, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 170
; LENGTH: 5312
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-170

Alignment Scores:

Pred. No.:	5.93e-11	Length:	5312
Score:	22.00	Matches:	22
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.54%	Indels:	0
DB:	14	Gaps:	0

US-09-787-097-12 (1-1429) x US-10-311-455-170 (1-5312)

```
Qy          47 LeuArgLeuProArgLeuLeuSerProProLeuArgProArgLeuLeuLeuLeuLeuLeu 66
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         174 CTACGCCTCCCGCGACTACTATCTCCACCGCTACGACCACGACTACTACTACTACTATTA 115

Qy          67 LeuLeu 68
               ||||||
Db         114 TTACTC 109
```

RESULT 37
US-09-908-975-15421
; Sequence 15421, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15421
; LENGTH: 60

```
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-908-975-15421

Alignment Scores:
Pred. No.:                8.43e-11        Length:         60
Score:                    20.00           Matches:        20
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              1.40%           Indels:         0
DB:                       10              Gaps:           0

US-09-787-097-12 (1-1429) x US-09-908-975-15421 (1-60)

Qy       1183 ValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsn 1202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTGGCTACTCCTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAAT 60

RESULT 38
US-09-908-975-29990
; Sequence 29990, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
;  APPLICANT: SHOSHAN, Avi
;  APPLICANT:  WASSERMAN, Alon
;  APPLICANT:  MINTZ, Eli
;  APPLICANT:  MINTZ, Liat
;  APPLICANT:  FAIGLER, Simchon
;  TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
;  TITLE OF INVENTION:  THAT POPULATE A TRANSCRIPTOME
;  FILE REFERENCE: 36688-0005
;  CURRENT APPLICATION NUMBER: US/09/908,975
;  CURRENT FILING DATE:  2001-07-20
;  PRIOR APPLICATION NUMBER: US 60/287,724
;  PRIOR FILING DATE: 2001-05-02
;  PRIOR APPLICATION NUMBER: US 60/221,607
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 32337
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 29990
;   LENGTH: 65
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-908-975-29990

Alignment Scores:
Pred. No.:                8.8e-09         Length:         65
Score:                    18.00           Matches:        18
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              1.26%           Indels:         0
DB:                       10              Gaps:           0

US-09-787-097-12 (1-1429) x US-09-908-975-29990 (1-65)

Qy        946 SerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAla 963
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         10 AGCTCGGAGTGCATGTGGTGCAGTAACATGAAGCAGTGTGTGGACTCCAATGCC 63

RESULT 39
US-09-728-446-450
; Sequence 450, Application US/09728446
; Patent No. US20020081668A1
; GENERAL INFORMATION:
;  APPLICANT: Friedrich, Glenn
;  APPLICANT:  Zambrowicz, Brian
;  APPLICANT:  Sands, Arthur T.
;  TITLE OF INVENTION: No. US20020081668A1el Murine Polynucleotide Sequences
;  TITLE OF INVENTION:  and Mutant Cells and Mutant Animals Defined Thereby
;  FILE REFERENCE: LEX-0101-USA
;  CURRENT APPLICATION NUMBER: US/09/728,446
;  CURRENT FILING DATE:  2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/168,270
;  PRIOR FILING DATE: 1999-12-01
;  NUMBER OF SEQ ID NOS: 1461
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 450
;   LENGTH: 233
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(233)
;   OTHER INFORMATION: n = A,T,C or G
US-09-728-446-450

Alignment Scores:
Pred. No.:                2.83e-06        Length:         233
Score:                    16.00           Matches:        16
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              1.12%           Indels:         0
DB:                       9               Gaps:           0

US-09-787-097-12 (1-1429) x US-09-728-446-450 (1-233)

Qy         62 LeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeuLeu 77
              |||||||||||||||||||||||||||||||||||||||||||||||
Db         60 CTCCTGCTGCTGCTGCTGCTGCCTCCGCCCCTGCTACTGCTGCTGTTG 107

RESULT 40
US-09-925-300-486/c
; Sequence 486, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
;  APPLICANT: Craig Rosen,
;  APPLICANT:  Steve Ruben
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA101
;  CURRENT APPLICATION NUMBER: US/09/925,300
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05988
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1890
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 486
;   LENGTH: 2572
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (823)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-486

Alignment Scores:
Pred. No.:                0.00263         Length:         2572
Score:                    14.00           Matches:        14
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              0.98%           Indels:         0
DB:                       9               Gaps:           0

US-09-787-097-12 (1-1429) x US-09-925-300-486 (1-2572)

Qy         64 LeuLeuLeuLeuProProProLeuLeuLeuLeuLeuLeu 77
              ||||||||||||||||||||||||||||||||||||||||
Db         60 CTGCTGCTGCTGCTGCCACCGCCGCTGCTGCTGCTGCTGCTG 19

RESULT 41
US-10-198-846-13651/c
; Sequence 13651, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
```

; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13651
; LENGTH: 3897
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3878, 3879, 3880, 3881, 3882, 3883, 3884, 3885, 3886, 3887,
; LOCATION: 3888, 3889, 3890, 3891, 3892, 3893, 3894, 3895, 3896, 3897
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-13651

Alignment Scores:

Pred. No.:	0.00389	Length:	3897
Score:	14.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.98%	Indels:	0
DB:	14	Gaps:	0

US-09-787-097-12 (1-1429) x US-10-198-846-13651 (1-3897)

```
Qy         64 LeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeuLeu 77
              |||||||||||||||||||||||||||||||||||||||||||
Db         80 CTGCTGCTGCTGCTGCCACCGCCGCTGCTGCTGCTGCTGCTG 39
```

RESULT 42
US-10-240-425-1425/c
; Sequence 1425, Application US/10240425
; Publication No. US20040033502A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Amanda
; APPLICANT: Boland, Joseph F.
; APPLICANT: Lord, Reginald V.
; APPLICANT: Alvarez, Chris
; APPLICANT: Wetzel, Jon C.
; APPLICANT: Scherf, Uwe
; APPLICANT: Vockley, Joseph G.
; TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue
; FILE REFERENCE: 44921-5026
; CURRENT APPLICATION NUMBER: US/10/240,425
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: PCT/US01/09847
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/193,446
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 1588
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1425
; LENGTH: 3901
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20040033502A1 U25997
US-10-240-425-1425

Alignment Scores:

Pred. No.:	0.00389	Length:	3901
Score:	14.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.98%	Indels:	0
DB:	12	Gaps:	0

US-09-787-097-12 (1-1429) x US-10-240-425-1425 (1-3901)

```
Qy         64 LeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeuLeu 77
              |||||||||||||||||||||||||||||||||||||||||||
Db         80 CTGCTGCTGCTGCTGCCACCGCCGCTGCTGCTGCTGCTGCTG 39
```

RESULT 43
US-10-177-293-440/c
; Sequence 440, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 440
; LENGTH: 3901
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-177-293-440

Alignment Scores:

Pred. No.:	0.00389	Length:	3901
Score:	14.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.98%	Indels:	0
DB:	14	Gaps:	0

US-09-787-097-12 (1-1429) x US-10-177-293-440 (1-3901)

```
Qy         64 LeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeuLeu 77
              |||||||||||||||||||||||||||||||||||||||||||
Db         80 CTGCTGCTGCTGCTGCCACCGCCGCTGCTGCTGCTGCTGCTG 39
```

RESULT 44

US-10-465-572-17/c
; Sequence 17, Application US/10465572
; Publication No. US20030207840A1
; GENERAL INFORMATION:
; APPLICANT: Riggins, Gregory
; APPLICANT: Lal, Anita
; TITLE OF INVENTION: GENES INDUCED BY HYPOXIA
; FILE REFERENCE: 000250.00012
; CURRENT APPLICATION NUMBER: US/10/465,572
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US/10/201,642
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/307,600
; PRIOR FILING DATE: 2001-07-26
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 3901
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-465-572-17

Alignment Scores:

Pred. No.:	0.00389	Length:	3901
Score:	14.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.98%	Indels:	0
DB:	15	Gaps:	0

US-09-787-097-12 (1-1429) x US-10-465-572-17 (1-3901)

```
Qy         64 LeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeuLeu 77
              |||||||||||||||||||||||||||||||||||||||||||
Db         80 CTGCTGCTGCTGCTGCCACCGCCGCTGCTGCTGCTGCTGCTG 39
```

RESULT 45
US-10-159-563-292/c
; Sequence 292, Application US/10159563
; Publication No. US20040009154A1
; GENERAL INFORMATION:
; APPLICANT: Khan, Javed
; APPLICANT: Ringner, Markus
; APPLICANT: Peterson, Carsten
; APPLICANT: Meltzer, Paul
; TITLE OF INVENTION: SELECTIONS OF GENES AND METHODS OF USING THE SAME FOR
; TITLE OF INVENTION: DIAGNOSIS AND FOR TARGETING THE THERAPY OF SELECT CANCERS
; FILE REFERENCE: 11613.56USI1
; CURRENT APPLICATION NUMBER: US/10/159,563
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 10/133,937
; PRIOR FILING DATE: 2002-04-25
; NUMBER OF SEQ ID NOS: 444
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 292
; LENGTH: 3901
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-159-563-292

Alignment Scores:

Pred. No.:	0.00389	Length:	3901
Score:	14.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.98%	Indels:	0
DB:	15	Gaps:	0

US-09-787-097-12 (1-1429) x US-10-159-563-292 (1-3901)

```
Qy         64 LeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeuLeu 77
              |||||||||||||||||||||||||||||||||||||||||||
Db         80 CTGCTGCTGCTGCTGCCACCGCCGCTGCTGCTGCTGCTGCTG 39
```

RESULT 46
US-10-372-683-40/c
; Sequence 40, Application US/10372683
; Publication No. US20040009171A1
; GENERAL INFORMATION:
; APPLICANT: GERRITSEN, MARY E.
; APPLICANT: PEALE JR., FRANKLIN V.
; APPLICANT: WU, THOMAS D.
; TITLE OF INVENTION: METHODS FOR THE TREATMENT OF CARCINOMA
; FILE REFERENCE: P1928R1P1
; CURRENT APPLICATION NUMBER: US/10/372,683
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 10/271,690
; PRIOR FILING DATE: 2002-10-16
; PRIOR APPLICATION NUMBER: US 60/344,534
; PRIOR FILING DATE: 2001-10-18
; NUMBER OF SEQ ID NOS: 49
; SEQ ID NO 40
; LENGTH: 3901
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-372-683-40

Alignment Scores:

Pred. No.:	0.00389	Length:	3901
Score:	14.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.98%	Indels:	0
DB:	15	Gaps:	0

US-09-787-097-12 (1-1429) x US-10-372-683-40 (1-3901)

```
Qy         64 LeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeuLeu 77
              |||||||||||||||||||||||||||||||||||||||||||
Db         80 CTGCTGCTGCTGCTGCCACCGCCGCTGCTGCTGCTGCTGCTG 39
```

RESULT 47
US-10-175-523-57/c
; Sequence 57, Application US/10175523
; Publication No. US20030096264A1
; GENERAL INFORMATION:
; APPLICANT: Brockman, Jeffrey
; APPLICANT: Evans, David
; APPLICANT: Hook, Derek
; APPLICANT: Klimczak, Leszek
; APPLICANT: Laeng, Pascal
; APPLICANT: Palfreyman, Michael
; APPLICANT: Rajan, Prithi
; TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)
; FILE REFERENCE: 3235/1J795-US3
; CURRENT APPLICATION NUMBER: US/10/175,523
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US 60/299,151
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/317,828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/325,150
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/333,047
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 60/349,936
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/361,834
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57
; LENGTH: 225883
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-175-523-57

```
Alignment Scores:
Pred. No.:              0.178          Length:         225883
Score:                  14.00          Matches:        14
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            0.98%          Indels:         0
DB:                     14             Gaps:           0

US-09-787-097-12 (1-1429) x US-10-175-523-57 (1-225883)

Qy         61 LeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeu 74
              ||||||||||||||||||||||||||||||||||||||||
Db      69681 CTCCTTCTTCTTCTTCTTCTTCCTCCTCCTCTTCTTCTT 69640

RESULT 48
US-10-029-386-14368
; Sequence 14368, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 14368
;   LENGTH: 105
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO CHR22_209.0
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.8
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.8
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.97
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
US-10-029-386-14368

Alignment Scores:
Pred. No.:              0.00127        Length:         105
Score:                  13.00          Matches:        13
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            0.91%          Indels:         0
DB:                     14             Gaps:           0

US-09-787-097-12 (1-1429) x US-10-029-386-14368 (1-105)

Qy         65 LeuLeuLeuLeuProProProLeuLeuLeuLeuLeuLeu 77
              |||||||||||||||||||||||||||||||||||||
Db          5 CTCCTCCTCCTTCCTCCTCCTCTCCTCCTCCTTCTTCTC 43

RESULT 49
US-10-027-632-181373/c
; Sequence 181373, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 181373
;   LENGTH: 520
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-181373

Alignment Scores:
Pred. No.:              0.00574        Length:         520
Score:                  13.00          Matches:        13
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            0.91%          Indels:         0
DB:                     15             Gaps:           0

US-09-787-097-12 (1-1429) x US-10-027-632-181373 (1-520)

Qy         65 LeuLeuLeuLeuProProProLeuLeuLeuLeuLeuLeu 77
              |||||||||||||||||||||||||||||||||||||
Db        130 CTACTATTACTACCACCACCACTACTACTGCTGCTGCTG 92

RESULT 50
US-10-029-386-664
; Sequence 664, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GE
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 664
;   LENGTH: 575
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO CHR22_209.0
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.8
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.8
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.97
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
;   OTHER INFORMATION: SWISSPROT HIT: P44087, EVALUE 8.40e-01
;   OTHER INFORMATION: EST_HUMAN HIT: BE877601.1, EVALUE 2.00e-01
;   OTHER INFORMATION: NT HIT: X93494.1, EVALUE 5.50e-01
US-10-029-386-664

Alignment Scores:
Pred. No.:              0.00632        Length:         575
Score:                  13.00          Matches:        13
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            0.91%          Indels:         0
DB:                     14             Gaps:           0
```

US-09-787-097-12 (1-1429) x US-10-029-386-664 (1-575)

```
Qy          65 LeuLeuLeuLeuProProProLeuLeuLeuLeuLeuLeu 77
               ||||||||||||||||||||||||||||||||||||||||
Db         395 CTCCTCCTCCTTCCTCCTCCTCTCCTCCTCCTTCTTCTC 433
```

RESULT 51
US-10-027-632-155757
; Sequence 155757, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 155757
; LENGTH: 935
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-155757

Alignment Scores:

Pred. No.:	0.00998	Length:	935
Score:	13.00	Matches:	13
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.91%	Indels:	0
DB:	15	Gaps:	0

US-09-787-097-12 (1-1429) x US-10-027-632-155757 (1-935)

```
Qy          61 LeuLeuLeuLeuLeuLeuLeuProProProLeuLeu 73
               |||||||||||||||||||||||||||||||||||||
Db         530 CTGCTGTTATTGCTACTGCTGCTGCCGCCGCCTCTGCTT 568
```

RESULT 52
US-10-108-260A-876
; Sequence 876, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 876
; LENGTH: 2257
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-876

Alignment Scores:

Pred. No.:	0.0229	Length:	2257
Score:	13.00	Matches:	13
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.91%	Indels:	0
DB:	15	Gaps:	0

US-09-787-097-12 (1-1429) x US-10-108-260A-876 (1-2257)

```
Qy          61 LeuLeuLeuLeuLeuLeuLeuProProProLeuLeu 73
               |||||||||||||||||||||||||||||||||||||
Db          64 CTGCTGTTATTGCTACTGCTGCTGCCGCCGCCTCTGCTT 102
```

RESULT 53
US-09-965-529-39
; Sequence 39, Application US/09965529
; Publication No. US20020182671A1
; GENERAL INFORMATION:
; APPLICANT: LAL, Preeti
; APPLICANT: YUE, Henry
; APPLICANT: TANG, Y. Tom
; APPLICANT: BANDMAN, Olga
; APPLICANT: BURFORD, Neil
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0731 USA
; CURRENT APPLICATION NUMBER: US/09/965,529
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/149,641; 60/164,203; PCT/US00/22315
; PRIOR FILING DATE: 1999-08-17; 1999-11-09; 2000-08-14
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PERL Program
; SEQ ID NO 39
; LENGTH: 2656
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020182671A1 997947CB1
; NAME/KEY: unsure
; LOCATION: 2554, 2587, 2604, 2606, 2611, 2644, 2646-2647, 2651
; OTHER INFORMATION: a, t, c, g, or other
US-09-965-529-39

Alignment Scores:

Pred. No.:	0.0267	Length:	2656
Score:	13.00	Matches:	13
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.91%	Indels:	0
DB:	9	Gaps:	0

US-09-787-097-12 (1-1429) x US-09-965-529-39 (1-2656)

```
Qy          61 LeuLeuLeuLeuLeuLeuLeuProProProLeuLeu 73
               |||||||||||||||||||||||||||||||||||||
Db          86 CTGCTGTTATTGCTACTGCTGCTGCCGCCGCCTCTGCTT 124
```

RESULT 54
US-09-969-680A-39
; Sequence 39, Application US/09969680A
; Publication No. US20030124649A1
; GENERAL INFORMATION:
; APPLICANT: LAL, Preeti; YUE, Henry
; APPLICANT: TANG, Y. Tom; BANDMAN, Olga
; APPLICANT: BURFORD, Neil; AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.; LU, Dyung Aina M.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0731-1 USA

```
;  CURRENT APPLICATION NUMBER: US/09/969,680A
;  CURRENT FILING DATE:  2001-10-02
;  PRIOR APPLICATION NUMBER: US00/22315
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/149,641
;  PRIOR FILING DATE: 1999-08-17
;  PRIOR APPLICATION NUMBER: 60/164,203
;  PRIOR FILING DATE: 1999-11-09
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: PERL Program
; SEQ ID NO 39
;   LENGTH: 2656
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20030124649A1 997947CB1
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1) ... (2656)
;   OTHER INFORMATION: a, t, c, g, or other
US-09-969-680A-39

Alignment Scores:
Pred. No.:                0.0267          Length:          2656
Score:                    13.00           Matches:         13
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              0.91%           Indels:          0
DB:                       10              Gaps:            0

US-09-787-097-12 (1-1429) x US-09-969-680A-39 (1-2656)

Qy          61 LeuLeuLeuLeuLeuLeuLeuProProProLeuLeu 73
               ||||||||||||||||||||||||||||||||||||||
Db          86 CTGCTGTTATTGCTACTGCTGCTGCCGCCGCCTCTGCTT 124

RESULT 55
US-10-301-822-3
; Sequence 3, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals, Inc.
;  APPLICANT:  Berger, Allison
;  APPLICANT:  Guillemette, Tracy L.
;  APPLICANT:  Kamatkar, Shubhangi
;  APPLICANT:  Schlegel, Robert
;  APPLICANT:  Monahan, John E.
;  APPLICANT:  Thibodeau, Stephen N.
;  APPLICANT:  Burgart, Lawrence J.
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
;  TITLE OF INVENTION:  METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
;  TITLE OF INVENTION:  THERAPY OF COLON CANCER
;  FILE REFERENCE: MPM01-029P2RNM
;  CURRENT APPLICATION NUMBER: US/10/301,822
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 60/339,971
;  PRIOR FILING DATE: 2001-12-10
;  PRIOR APPLICATION NUMBER: US 60/361,978
;  PRIOR FILING DATE: 2002-03-05
;  PRIOR APPLICATION NUMBER: US 60/381,988
;  PRIOR FILING DATE: 2002-05-20
;  NUMBER OF SEQ ID NOS: 228
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 6691
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (82)...(3717)
US-10-301-822-3

Alignment Scores:
Pred. No.:                0.0637          Length:          6691
Score:                    13.00           Matches:         13
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              0.91%           Indels:          0
DB:                       14              Gaps:            0

US-09-787-097-12 (1-1429) x US-10-301-822-3 (1-6691)

Qy          61 LeuLeuLeuLeuLeuLeuLeuProProProLeuLeu 73
               ||||||||||||||||||||||||||||||||||||||
Db         127 CTGCTGCTGCTGCTGCTGCTGCCGCCGCCGCTCCTG 165

RESULT 56
US-09-788-711A-3
; Sequence 3, Application US/09788711A
; Patent No. US20020058328A1
; GENERAL INFORMATION:
;  APPLICANT: Tania Tamsin Testa
;  TITLE OF INVENTION: NOVEL COMPOUNDS
;  FILE REFERENCE: GP-30225
;  CURRENT APPLICATION NUMBER: US/09/788,711A
;  CURRENT FILING DATE:  2001-02-20
;  PRIOR APPLICATION NUMBER: 0004196.2
;  PRIOR FILING DATE: 2000-02-19
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 8772
;   TYPE: DNA
;   ORGANISM: HOMO SAPIENS
US-09-788-711A-3

Alignment Scores:
Pred. No.:                0.0821          Length:          8772
Score:                    13.00           Matches:         13
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              0.91%           Indels:          0
DB:                       9               Gaps:            0

US-09-787-097-12 (1-1429) x US-09-788-711A-3 (1-8772)

Qy          61 LeuLeuLeuLeuLeuLeuLeuProProProLeuLeu 73
               ||||||||||||||||||||||||||||||||||||||
Db          52 CTGCTGCTGTTGCTGCTGCTGCTGCCGCCGCCACTATTG 90

RESULT 57
US-09-788-711A-1
; Sequence 1, Application US/09788711A
; Patent No. US20020058328A1
; GENERAL INFORMATION:
;  APPLICANT: Tania Tamsin Testa
;  TITLE OF INVENTION: NOVEL COMPOUNDS
;  FILE REFERENCE: GP-30225
;  CURRENT APPLICATION NUMBER: US/09/788,711A
;  CURRENT FILING DATE:  2001-02-20
;  PRIOR APPLICATION NUMBER: 0004196.2
;  PRIOR FILING DATE: 2000-02-19
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 8871
;   TYPE: DNA
;   ORGANISM: HOMO SAPIENS
US-09-788-711A-1

Alignment Scores:
Pred. No.:                0.083           Length:          8871
Score:                    13.00           Matches:         13
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
```

```
Query Match:            0.91%       Indels:        0
DB:                     9           Gaps:          0

US-09-787-097-12 (1-1429) x US-09-788-711A-1 (1-8871)

Qy         61 LeuLeuLeuLeuLeuLeuLeuProProProLeuLeu 73
              |||||||||||||||||||||||||||||||||||||
Db         52 CTGCTGCTGTTGCTGCTGCTGCCGCCGCCACTATTG 90

RESULT 58
US-10-311-623-21
; Sequence 21, Application US/10311623
; Publication No. US20040023244A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE GENOMICS, INC.; GRIFFIN, Jennifer A.
;  APPLICANT:  KALLICK, Deborah A.; TRIBOULEY, Catherine M.
;  APPLICANT:  YUE, Henry; NGUYEN, Danniel B.
;  APPLICANT:  TANG, Y. Tom; LAL, Preeti G.
;  APPLICANT:  POLICKY, Jennifer L.; AZIMZAI, Yalda
;  APPLICANT:  LU, Dyung Aina M.; GRAUL, Richard C.
;  APPLICANT:  YAO, Monique G.; BURFORD, Neil
;  APPLICANT:  HAFALIA, April J. A.; BAUGHN, Mariah R.
;  APPLICANT:  BANDMAN, Olga; ARVIZU, Chandra S.
;  APPLICANT:  YANG, Junming; XU, Yuming
;  APPLICANT:  GANDHI, Ameena R.; WARREN, Bridget A.
;  APPLICANT:  DING, Li; SANJANWALA, Madhusudan M.
;  APPLICANT:  DUGGAN, Brendan M.; LU, Yan
;  TITLE OF INVENTION: RECEPTORS
;  FILE REFERENCE: PF-0793 USN
;  CURRENT APPLICATION NUMBER: US/10/311,623
;  CURRENT FILING DATE:  2002-12-17
;  PRIOR APPLICATION NUMBER: US 01/19942
;  PRIOR FILING DATE: 2001-06-21
;  PRIOR APPLICATION NUMBER: US 60/214,027
;  PRIOR FILING DATE: 2000-06-21
;  PRIOR APPLICATION NUMBER: US 60/228,045
;  PRIOR FILING DATE: 2000-08-25
;  PRIOR APPLICATION NUMBER: US 60/255,104
;  PRIOR FILING DATE: 2000-12-12
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PERL Program
; SEQ ID NO 21
;   LENGTH: 9321
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20040023244A1 6977010CB1
US-10-311-623-21

Alignment Scores:
Pred. No.:              0.087       Length:        9321
Score:                  13.00       Matches:       13
Percent Similarity:     100.00%     Conservative:  0
Best Local Similarity:  100.00%     Mismatches:    0
Query Match:            0.91%       Indels:        0
DB:                     16          Gaps:          0

US-09-787-097-12 (1-1429) x US-10-311-623-21 (1-9321)

Qy         61 LeuLeuLeuLeuLeuLeuLeuProProProLeuLeu 73
              |||||||||||||||||||||||||||||||||||||
Db        277 CTGCTGCTGTTGCTGCTGCTGCCGCCGCCACTATTG 315

RESULT 59
US-10-225-567A-523
; Sequence 523, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
;  APPLICANT: LifeSpan Biosciences
;  APPLICANT:  Brown, Joseph P.
;  APPLICANT:  Burmer, Glenna C.
;  APPLICANT:  Roush, Christine L.
;  TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
;  FILE REFERENCE: 1920-4-4
;  CURRENT APPLICATION NUMBER: US/10/225,567A
;  CURRENT FILING DATE:  2001-12-19
;  PRIOR APPLICATION NUMBER: 60/257,144
;  PRIOR FILING DATE: 2000-12-19
;  NUMBER OF SEQ ID NOS: 2292
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 523
;   LENGTH: 10531
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-225-567A-523

Alignment Scores:
Pred. No.:              0.0976      Length:        10531
Score:                  13.00       Matches:       13
Percent Similarity:     100.00%     Conservative:  0
Best Local Similarity:  100.00%     Mismatches:    0
Query Match:            0.91%       Indels:        0
DB:                     14          Gaps:          0

US-09-787-097-12 (1-1429) x US-10-225-567A-523 (1-10531)

Qy         61 LeuLeuLeuLeuLeuLeuLeuProProProLeuLeu 73
              |||||||||||||||||||||||||||||||||||||
Db        114 CTGCTGCTGTTGCTGCTGCTGCCGCCGCCACTATTG 152

RESULT 60
US-10-292-798-931
; Sequence 931, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
;  APPLICANT: SUWA, MAKIKO
;  APPLICANT:  ASAI, KIYOSHI
;  APPLICANT:  AKIYAMA, YUTAKA
;  APPLICANT:  ABURATANI, HIROYUKI
;  TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
;  FILE REFERENCE: 084335/166
;  CURRENT APPLICATION NUMBER: US/10/292,798
;  CURRENT FILING DATE:  2002-11-13
;  PRIOR APPLICATION NUMBER: 10/017,161
;  PRIOR FILING DATE: 2001-12-18
;  PRIOR APPLICATION NUMBER: JP 2001-246789
;  PRIOR FILING DATE: 2001-06-18
;  NUMBER OF SEQ ID NOS: 2070
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 931
;   LENGTH: 24370
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   LOCATION: source
;   FEATURE:
;   LOCATION: (1)..(24370)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (201)..(3510)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (8553)..(9200)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (11163)..(11385)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (11634)..(11749)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (11929)..(12017)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (12408)..(12565)
```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (12927)..(13090)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (13263)..(13388)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (13732)..(13898)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (14199)..(14486)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (14575)..(14749)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (14989)..(15130)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (15312)..(15432)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (15856)..(16039)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (16226)..(16327)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (17669)..(17763)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (17972)..(18181)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (18702)..(18885)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (19001)..(19129)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (19231)..(19411)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (19545)..(19712)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (19815)..(19941)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (20054)..(20254)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (20547)..(20721)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (21048)..(21160)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (21338)..(21417)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (21507)..(21633)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (21721)..(21843)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (22399)..(22526)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (22761)..(22849)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (22954)..(23148)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (23287)..(23457)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (23557)..(23791)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (24143)..(24170)
US-10-292-798-931

Alignment Scores:

Pred. No.:	0.215	Length:	24370
Score:	13.00	Matches:	13
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.91%	Indels:	0
DB:	15	Gaps:	0

US-09-787-097-12 (1-1429) x US-10-292-798-931 (1-24370)

```
Qy         61 LeuLeuLeuLeuLeuLeuLeuProProProLeuLeu 73
              |||||||||||||||||||||||||||||||||||||
Db        252 CTGCTGCTGTTGCTGCTGCTGCCGCCGCCACTATTG 290
```

RESULT 61
US-10-034-650-31
; Sequence 31, Application US/10034650
; Publication No. US20030216558A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David
; APPLICANT: Engelhard, Eric
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000128
; CURRENT APPLICATION NUMBER: US/10/034,650
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 09/474,377
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 68233
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-034-650-31

Alignment Scores:

Pred. No.:	0.567	Length:	68233
Score:	13.00	Matches:	13
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.91%	Indels:	0
DB:	15	Gaps:	0

US-09-787-097-12 (1-1429) x US-10-034-650-31 (1-68233)

```
Qy         61 LeuLeuLeuLeuLeuLeuLeuProProProLeuLeu 73
              |||||||||||||||||||||||||||||||||||||
Db      65999 CTCCTTCTTCTTCTTCTTCTTCCTCCTCCTCTCCTC 66037
```

RESULT 62
US-09-864-761-24815/c
; Sequence 24815, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 24815
; LENGTH: 287
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC017016.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 8.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.1
US-09-864-761-24815

Alignment Scores:

Pred. No.:	0.0323	Length:	287
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.84%	Indels:	0
DB:	9	Gaps:	0

US-09-787-097-12 (1-1429) x US-09-864-761-24815 (1-287)

```
Qy         65 LeuLeuLeuLeuProProProLeuLeuLeuLeuLeu 76
              ||||||||||||||||||||||||||||||||||||||
Db        167 CTTCTTCTTCTTCCTCCTCCTCTTCTTCTCCTTCTC 132
```

RESULT 63
US-10-027-632-28416/c
; Sequence 28416, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 28416
; LENGTH: 361
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-28416

Alignment Scores:

Pred. No.:	0.0401	Length:	361
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.84%	Indels:	0
DB:	15	Gaps:	0

US-09-787-097-12 (1-1429) x US-10-027-632-28416 (1-361)

```
Qy         66 LeuLeuLeuProProProLeuLeuLeuLeuLeu 77
              |||||||||||||||||||||||||||||||||||
Db        325 CTCCTCCTTCCTCCTCCCCTTCTTCTTCTTCTT 290
```

RESULT 64
US-10-198-846-13018/c
; Sequence 13018, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13018
; LENGTH: 394
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-198-846-13018

```
Alignment Scores:
Pred. No.:                0.0435          Length:          394
Score:                    12.00           Matches:         12
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              0.84%           Indels:          0
DB:                       14              Gaps:            0

US-09-787-097-12 (1-1429) x US-10-198-846-13018 (1-394)

Qy          61 LeuLeuLeuLeuLeuLeuLeuProProProLeu 72
               ||||||||||||||||||||||||||||||||||
Db          71 CTCCTCCTCCTCCTCCTGCTGCCTCCCCCGTTA 36

RESULT 65
US-10-040-739-599
; Sequence 599, Application US/10040739
; Publication No. US20020173635A1
;   GENERAL INFORMATION:
;        APPLICANT: Jacobs, Kenneth
;                   McCoy, John
;                   LaVallie, Edward
;                   Racie, Lisa
;                   Merberg, David
;                   Treacy, Maurice
;                   Spaulding, Vikki
;        TITLE OF INVENTION: SECRETED, EXPRESSED SEQUENCE TAGS
;        NUMBER OF SEQUENCES: 1519
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Genetics Institute, Inc.
;             STREET: 87 CambridgePark Drive
;             CITY: Cambridge
;             STATE: Massachusetts
;             COUNTRY: U.S.A
;             ZIP: 02140
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy Disk
;             COMPUTER: IBM PC Compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/040,739
;             FILING DATE: 07-Jan-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/036,520
;             FILING DATE: 03-JUN-1998
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Brown, Scott A.
;             REGISTRATION NUMBER: 32,724
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617) 498-8224
;             TELEFAX: (617) 876-5851
;   INFORMATION FOR SEQ ID NO: 599:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 411 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 599:
US-10-040-739-599

Alignment Scores:
Pred. No.:                0.0453          Length:          411
Score:                    12.00           Matches:         12
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              0.84%           Indels:          0
DB:                       13              Gaps:            0

US-09-787-097-12 (1-1429) x US-10-040-739-599 (1-411)

Qy          60 ArgLeuLeuLeuLeuLeuLeuLeuProProPro 71
               |||||||||||||||||||||||||||||||||
Db         118 AGATTGCTGCTGCTGCTGCTGCTGCCGCCGCCG 153

RESULT 66
US-09-864-761-2885
; Sequence 2885, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 2885
;   LENGTH: 436
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AP000018.2
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.9
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4
US-09-864-761-2885
```

Alignment Scores:

Pred. No.:	0.0479	Length:	436
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.84%	Indels:	0
DB:	9	Gaps:	0

US-09-787-097-12 (1-1429) x US-09-864-761-2885 (1-436)

```
Qy         65 LeuLeuLeuLeuProProProLeuLeuLeuLeuLeu 76
              |||||||||||||||||||||||||||||||||||||
Db        335 TTACTACTACTACCACCACCATTACTACTACTGCTG 370
```

RESULT 67
US-10-198-846-5901/c
; Sequence 5901, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5901
; LENGTH: 455
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 81, 84, 188, 436, 445
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-5901

Alignment Scores:

Pred. No.:	0.0498	Length:	455
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.84%	Indels:	0
DB:	14	Gaps:	0

US-09-787-097-12 (1-1429) x US-10-198-846-5901 (1-455)

```
Qy         61 LeuLeuLeuLeuLeuLeuLeuLeuProProProLeu 72
              ||||||||||||||||||||||||||||||||||||
Db        130 CTCCTCCTCCTCCTCCTGCTGCCTCCCCCGTTA 95
```

RESULT 68
US-09-930-213-151/c
; Sequence 151, Application US/09930213
; Publication No. US20030170625A1
; GENERAL INFORMATION:
; APPLICANT: ROSENTHAL, ANDRE
; APPLICANT: HINZMANN, BERND
; APPLICANT: SCHAFER, REINHARD
; APPLICANT: ZUBER, JOHANNES
; APPLICANT: TCHE-NITSE, OLEG
; APPLICANT: GRIPS, MARTIN
; APPLICANT: HELLNEGEL, MARTIN
; APPLICANT: SCHMITZ, ANNE-CHANTAL
; APPLICANT: SERS, CHRISTINE
; TITLE OF INVENTION: DETECTION OF DIFFERENTIAL GENE EXPRESSIONS
; FILE REFERENCE: ALBRE-14
; CURRENT APPLICATION NUMBER: US/09/930,213
; CURRENT FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: DE 10004102.7
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 885
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 151
; LENGTH: 465
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-930-213-151

Alignment Scores:

Pred. No.:	0.0509	Length:	465
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.84%	Indels:	0
DB:	10	Gaps:	0

US-09-787-097-12 (1-1429) x US-09-930-213-151 (1-465)

```
Qy         60 ArgLeuLeuLeuLeuLeuLeuLeuProProPro 71
              |||||||||||||||||||||||||||||||||||
Db        226 AGATTGCTGCTGCTGCTGCTGCTGCCGCCGCCG 191
```

RESULT 69
US-10-027-632-108988/c
; Sequence 108988, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 108988
; LENGTH: 468
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(468)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-108988

Alignment Scores:

Pred. No.:	0.0512	Length:	468
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.84%	Indels:	0
DB:	15	Gaps:	0

US-09-787-097-12 (1-1429) x US-10-027-632-108988 (1-468)

```
Qy       1140 CysGluValGluAsnArgTyrGlnGlyAsnProLeu 1151
              ||||||||||||||||||||||||||||||||||||||
Db        413 TGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTC 378

RESULT 70
US-10-029-386-11510/c
; Sequence 11510, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 11510
;   LENGTH: 526
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC004690.1
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5
;   OTHER INFORMATION: EST_HUMAN HIT: AA065301.1, EVALUE 2.00e-81
;   OTHER INFORMATION: SWISSPROT HIT: P36814, EVALUE 2.80e+00
;   OTHER INFORMATION: NT HIT: gi9845300, EVALUE 7.80e+00
US-10-029-386-11510

Alignment Scores:
Pred. No.:                0.0571          Length:          526
Score:                    12.00           Matches:         12
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              0.84%           Indels:          0
DB:                       14              Gaps:            0

US-09-787-097-12 (1-1429) x US-10-029-386-11510 (1-526)

Qy          61 LeuLeuLeuLeuLeuLeuLeuProProProLeu 72
               ||||||||||||||||||||||||||||||||||||
Db         187 CTGCTGCTGCTGCTGTTGCTGCTGCCGCCGCCTCTG 152

RESULT 71
US-09-864-761-8070/c
; Sequence 8070, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 8070
;   LENGTH: 575
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC017016.3
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.5
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.7
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 8.1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.3
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.5
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.1
US-09-864-761-8070

Alignment Scores:
Pred. No.:                0.0621          Length:          575
Score:                    12.00           Matches:         12
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              0.84%           Indels:          0
DB:                       9               Gaps:            0

US-09-787-097-12 (1-1429) x US-09-864-761-8070 (1-575)

Qy          65 LeuLeuLeuLeuProProProLeuLeuLeuLeuLeu 76
               ||||||||||||||||||||||||||||||||||||
Db         220 CTTCTTCTTCTTCCTCCTCCTCTTCTTCTCCTTCTC 185

RESULT 72
US-10-101-464A-290
; Sequence 290, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
;  APPLICANT: Strabala, Timothy
;  APPLICANT:  Nieuwenhuizen, Nicolaas
;  APPLICANT:  Higgins, Colleen M.
;  TITLE OF INVENTION: Compositions Isolated from Plant Cells
;  TITLE OF INVENTION:  and Their Use in the Modification of Plant Cell Signaling
;  FILE REFERENCE: 11000.1020c2
;  CURRENT APPLICATION NUMBER: US/10/101,464A
;  CURRENT FILING DATE:  2002-03-18
;  PRIOR APPLICATION NUMBER: 09/704,302
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: 09/228,986
;  PRIOR FILING DATE: 1999-01-12
;  PRIOR APPLICATION NUMBER: 60/162,866
```

; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 290
; LENGTH: 641
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-10-101-464A-290

Alignment Scores:

Pred. No.:	0.0688	Length:	641
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.84%	Indels:	0
DB:	14	Gaps:	0

US-09-787-097-12 (1-1429) x US-10-101-464A-290 (1-641)

```
Qy         61 LeuLeuLeuLeuLeuLeuLeuProProProLeu 72
              ||||||||||||||||||||||||||||||||||
Db        324 CTTCTGCTTCTCCTCCTCCTTCTGCCTCCTCCTCTT 359
```

RESULT 73
US-10-029-386-20268
; Sequence 20268, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 20268
; LENGTH: 659
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005317.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.7
; OTHER INFORMATION: NT HIT: gi15300936, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P54098, EVALUE 4.00e-78
; OTHER INFORMATION: EST_HUMAN HIT: AL528934.1, EVALUE 0.00e+00
US-10-029-386-20268

Alignment Scores:

Pred. No.:	0.0706	Length:	659
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.84%	Indels:	0
DB:	14	Gaps:	0

US-09-787-097-12 (1-1429) x US-10-029-386-20268 (1-659)

```
Qy         61 LeuLeuLeuLeuLeuLeuLeuProProProLeu 72
              ||||||||||||||||||||||||||||||||||
Db        510 CTGCTGCTGCTGCTGCTGCTGCCGCCGCCGCTG 545
```

RESULT 74
US-10-424-599-17604/c
; Sequence 17604, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 17604
; LENGTH: 834
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_115900C.1
US-10-424-599-17604

Alignment Scores:

Pred. No.:	0.0882	Length:	834
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.84%	Indels:	0
DB:	12	Gaps:	0

US-09-787-097-12 (1-1429) x US-10-424-599-17604 (1-834)

```
Qy         61 LeuLeuLeuLeuLeuLeuLeuProProProLeu 72
              ||||||||||||||||||||||||||||||||||
Db        695 TTACTGTTGTTGTTATTGTTACTGCCACCACCACTA 660
```

RESULT 75
US-10-425-114-35604/c
; Sequence 35604, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 35604
; LENGTH: 1256
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMROB73064G02_FLI
US-10-425-114-35604

Alignment Scores:

Pred. No.:	0.13	Length:	1256
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.84%	Indels:	0
DB:	12	Gaps:	0

US-09-787-097-12 (1-1429) x US-10-425-114-35604 (1-1256)

```
Qy         58 ArgProArgLeuLeuLeuLeuLeuLeuLeuPro 69
```

```
                 ||||||||||||||||||||||||||||||||||||
Db        174 CGGCCCCGCCTGCTGCTGCTGCTGCTGCTGCCT 139
```

Search completed: March 1, 2004, 13:27:29
Job time : 1775 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 1, 2004, 07:13:00 ; Search time 236 Seconds
(without alignments)
3360.275 Million cell updates/sec

Title: US-09-787-097-12
Perfect score: 1429
Sequence: 1 MVAAAAATEARLRRRTAATA.....KSGAVRNRKQQPPAQPGTCI 1429

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1360453

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09787097/runat_26022004_063015_23781/app_query.fasta_1.1607
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=oligo.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=100 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=75 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09787097_@CGN_1_1_118_@runat_26022004_063015_23781 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1141	79.8	8589	3	US-09-245-041-14	Sequence 14, Appl
	2	979	68.5	4072	3	US-09-245-041-16	Sequence 16, Appl
	3	503	35.2	6370	3	US-09-245-041-12	Sequence 12, Appl
	4	496	34.7	2625	3	US-09-245-041-18	Sequence 18, Appl
	5	206	14.4	8827	3	US-09-245-041-1	Sequence 1, Appli
	6	94	6.6	2419	3	US-09-245-041-8	Sequence 8, Appli
c	7	82	5.7	246	4	US-09-702-705-966	Sequence 966, App
c	8	82	5.7	246	4	US-09-736-457-966	Sequence 966, App
c	9	82	5.7	246	4	US-09-614-124B-966	Sequence 966, App
c	10	82	5.7	246	4	US-09-671-325-966	Sequence 966, App
	11	76	5.3	1051	3	US-09-245-041-10	Sequence 10, Appl
	12	73	5.1	90050	3	US-09-245-041-5	Sequence 5, Appli
c	13	51	3.6	226	3	US-09-245-041-6	Sequence 6, Appli
c	14	45	3.1	221	3	US-09-245-041-7	Sequence 7, Appli
	15	33	2.3	5973	3	US-09-245-041-4	Sequence 4, Appli
c	16	30	2.1	17056	3	US-09-245-041-3	Sequence 3, Appli
	17	22	1.5	415	4	US-09-833-381-1916	Sequence 1916, Ap
	18	13	0.9	429	3	US-08-905-223-61	Sequence 61, Appl
	19	13	0.9	2023	4	US-09-491-522-6	Sequence 6, Appli
	20	13	0.9	2450	4	US-09-491-522-2	Sequence 2, Appli
	21	13	0.9	6692	4	US-09-491-522-1	Sequence 1, Appli
	22	12	0.8	1103	3	US-08-927-219-54	Sequence 54, Appl
	23	12	0.8	2426	3	US-09-028-327-2	Sequence 2, Appli
	24	12	0.8	2426	4	US-09-571-078A-2	Sequence 2, Appli
	25	11	0.8	639	4	US-09-336-536-57	Sequence 57, Appl
c	26	11	0.8	1432	4	US-09-620-312D-420	Sequence 420, App
	27	11	0.8	1739	4	US-09-148-545-11	Sequence 11, Appl
	28	11	0.8	1858	4	US-09-336-536-56	Sequence 56, Appl
	29	11	0.8	1899	4	US-09-919-060-15	Sequence 15, Appl
c	30	11	0.8	1899	4	US-09-919-060-16	Sequence 16, Appl
	31	11	0.8	2362	4	US-09-907-794A-194	Sequence 194, App
	32	11	0.8	2362	4	US-09-905-125A-194	Sequence 194, App
	33	11	0.8	2362	4	US-09-902-775A-194	Sequence 194, App
c	34	11	0.8	2520	1	US-08-405-254-9	Sequence 9, Appli
	35	11	0.8	2693	4	US-09-919-060-12	Sequence 12, Appl
c	36	11	0.8	2693	4	US-09-919-060-14	Sequence 14, Appl
c	37	11	0.8	2784	1	US-08-261-465-3	Sequence 3, Appli
c	38	11	0.8	2784	1	US-08-405-254-7	Sequence 7, Appli
c	39	11	0.8	3110	4	US-09-023-655-1072	Sequence 1072, Ap
c	40	11	0.8	3132	2	US-08-224-482-3	Sequence 3, Appli
c	41	11	0.8	3132	3	US-09-205-921-1	Sequence 1, Appli
c	42	11	0.8	3132	4	US-09-300-958A-32	Sequence 32, Appl
c	43	11	0.8	3132	4	US-09-702-705-320	Sequence 320, App
c	44	11	0.8	3132	4	US-09-736-457-320	Sequence 320, App
c	45	11	0.8	3132	4	US-09-614-124B-320	Sequence 320, App
c	46	11	0.8	3132	4	US-09-671-325-320	Sequence 320, App
c	47	11	0.8	3132	4	US-09-589-184-320	Sequence 320, App
	48	11	0.8	3200	4	US-08-884-569A-1	Sequence 1, Appli
c	49	11	0.8	3311	4	US-09-367-891A-5	Sequence 5, Appli
	50	11	0.8	4169	4	US-09-166-350-32	Sequence 32, Appl
c	51	11	0.8	5828	4	US-09-023-655-1386	Sequence 1386, Ap
	52	11	0.8	392000	4	US-10-027-983-11	Sequence 11, Appl
c	53	10	0.7	100	2	US-08-811-492-132	Sequence 132, App
c	54	10	0.7	109	2	US-08-811-492-133	Sequence 133, App
c	55	10	0.7	122	1	US-08-478-312-55	Sequence 55, Appl
c	56	10	0.7	122	1	US-08-485-302-55	Sequence 55, Appl
c	57	10	0.7	122	1	US-08-476-169-60	Sequence 60, Appl
c	58	10	0.7	122	1	US-08-484-083-60	Sequence 60, Appl
c	59	10	0.7	122	2	US-08-764-640-237	Sequence 237, App
c	60	10	0.7	122	3	US-08-973-225-225	Sequence 225, App
c	61	10	0.7	122	3	US-09-244-298A-237	Sequence 237, App
c	62	10	0.7	122	3	US-09-516-704-237	Sequence 237, App
c	63	10	0.7	122	4	US-09-549-090-225	Sequence 225, App
c	64	10	0.7	122	4	US-09-832-230A-237	Sequence 237, App
	65	10	0.7	577	3	US-08-938-548B-21	Sequence 21, Appl
	66	10	0.7	577	4	US-08-939-093A-21	Sequence 21, Appl
	67	10	0.7	577	4	US-09-211-823C-21	Sequence 21, Appl
c	68	10	0.7	614	4	US-09-325-932A-31	Sequence 31, Appl
	69	10	0.7	867	4	US-09-833-381-782	Sequence 782, App
	70	10	0.7	999	4	US-08-811-481-26	Sequence 26, Appl
	71	10	0.7	999	4	US-09-876-527-26	Sequence 26, Appl
	72	10	0.7	1011	4	US-08-811-481-27	Sequence 27, Appl
	73	10	0.7	1011	4	US-09-876-527-27	Sequence 27, Appl
c	74	10	0.7	1146	4	US-09-489-039A-3636	Sequence 3636, Ap
c	75	10	0.7	1276	3	US-09-177-325-2	Sequence 2, Appli
c	76	10	0.7	1276	3	US-09-411-812A-2	Sequence 2, Appli
c	77	10	0.7	1276	4	US-09-590-113-2	Sequence 2, Appli
c	78	10	0.7	1425	1	US-08-464-148-1	Sequence 1, Appli
c	79	10	0.7	1425	1	US-08-385-500-1	Sequence 1, Appli
c	80	10	0.7	1425	1	US-08-846-784-1	Sequence 1, Appli
c	81	10	0.7	1436	4	US-09-471-396-2	Sequence 2, Appli
c	82	10	0.7	1436	4	US-09-788-345-11	Sequence 11, Appl
	83	10	0.7	1537	2	US-08-839-008-1	Sequence 1, Appli
c	84	10	0.7	1578	4	US-09-130-491-10	Sequence 10, Appl
c	85	10	0.7	1768	4	US-09-833-381-523	Sequence 523, App

86	10	0.7	1885	4	US-09-023-655-1162	Sequence 1162, Ap
87	10	0.7	1970	3	US-08-938-548B-1	Sequence 1, Appli
88	10	0.7	1970	4	US-08-939-093A-1	Sequence 1, Appli
89	10	0.7	1970	4	US-09-211-823C-1	Sequence 1, Appli
c 90	10	0.7	2073	3	US-08-816-977-46	Sequence 46, Appl
c 91	10	0.7	2073	4	US-09-334-477-46	Sequence 46, Appl
c 92	10	0.7	2085	3	US-08-816-977-48	Sequence 48, Appl
c 93	10	0.7	2085	4	US-09-334-477-48	Sequence 48, Appl
c 94	10	0.7	2127	3	US-08-816-977-32	Sequence 32, Appl
c 95	10	0.7	2127	4	US-09-334-477-32	Sequence 32, Appl
c 96	10	0.7	2136	3	US-08-816-977-34	Sequence 34, Appl
c 97	10	0.7	2136	4	US-09-334-477-34	Sequence 34, Appl
98	10	0.7	2329	1	US-08-455-559-9	Sequence 9, Appli
99	10	0.7	2329	3	US-09-145-060-9	Sequence 9, Appli
100	10	0.7	2329	5	PCT-US94-00657-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1
US-09-245-041-14
; Sequence 14, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 8589
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-245-041-14

Alignment Scores:

Pred. No.:	0	Length:	8589
Score:	1141.00	Matches:	1341
Percent Similarity:	99.70%	Conservative:	0
Best Local Similarity:	99.70%	Mismatches:	2
Query Match:	79.85%	Indels:	4
DB:	3	Gaps:	0

US-09-787-097-12 (1-1429) x US-09-245-041-14 (1-8589)

```
Qy        87 AlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGluCysAspArgProCysVal 106
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       256 GCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAAGGAATGTGACCGGCCGTGTGTC 315

Qy       107 AsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTrpValGly 126
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       316 AACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGC 375

Qy       127 GluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPheValThr 146
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       376 GAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACA 435

Qy       147 AspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeuIleGluGlyGlnPro 166
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       436 GATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCA 495

Qy       167 AsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLeu 186
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       496 AATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCATTTA 555

Qy       187 TyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaAlaPheSerGlyLeuIle 206
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       556 TATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATT 615

Qy       207 ValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGlyTyrAla 226
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       616 GTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCC 675

Qy       227 LeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThrTyrSer 246
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       676 TTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTACAGT 735

Qy       247 PheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSerSer 266
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       736 TTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGC 795

Qy       267 Glu-ThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIleProHi 286
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       796 GA-TACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCA 854

Qy       286 sCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspValArgGl 306
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       855 CTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGG 914

Qy       306 yCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAlaAsnGl 326
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       915 ATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCA 974

Qy       326 nSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHisLysAl 346
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       975 GTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGC 1034

Qy       346 aValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHisSerAspTy 366
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1035 TGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTA 1094

Qy       366 rAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSerVa 386
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1095 TAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGT 1154

Qy       386 lAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyrMe 406
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1155 GAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACAT 1214

Qy       406 tTyrGlyGlyLysIleAspPro-ThrGlyAsnValThrAsnGluLeuArgValPheHisI 426
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db      1215 GTATGGAGGAAAAATTGATTC-AACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACA 1273

Qy       426 leHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaValValG 446
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1274 TTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTG 1333

Qy       446 lyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValIlePheG 466
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1334 GGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTG 1393

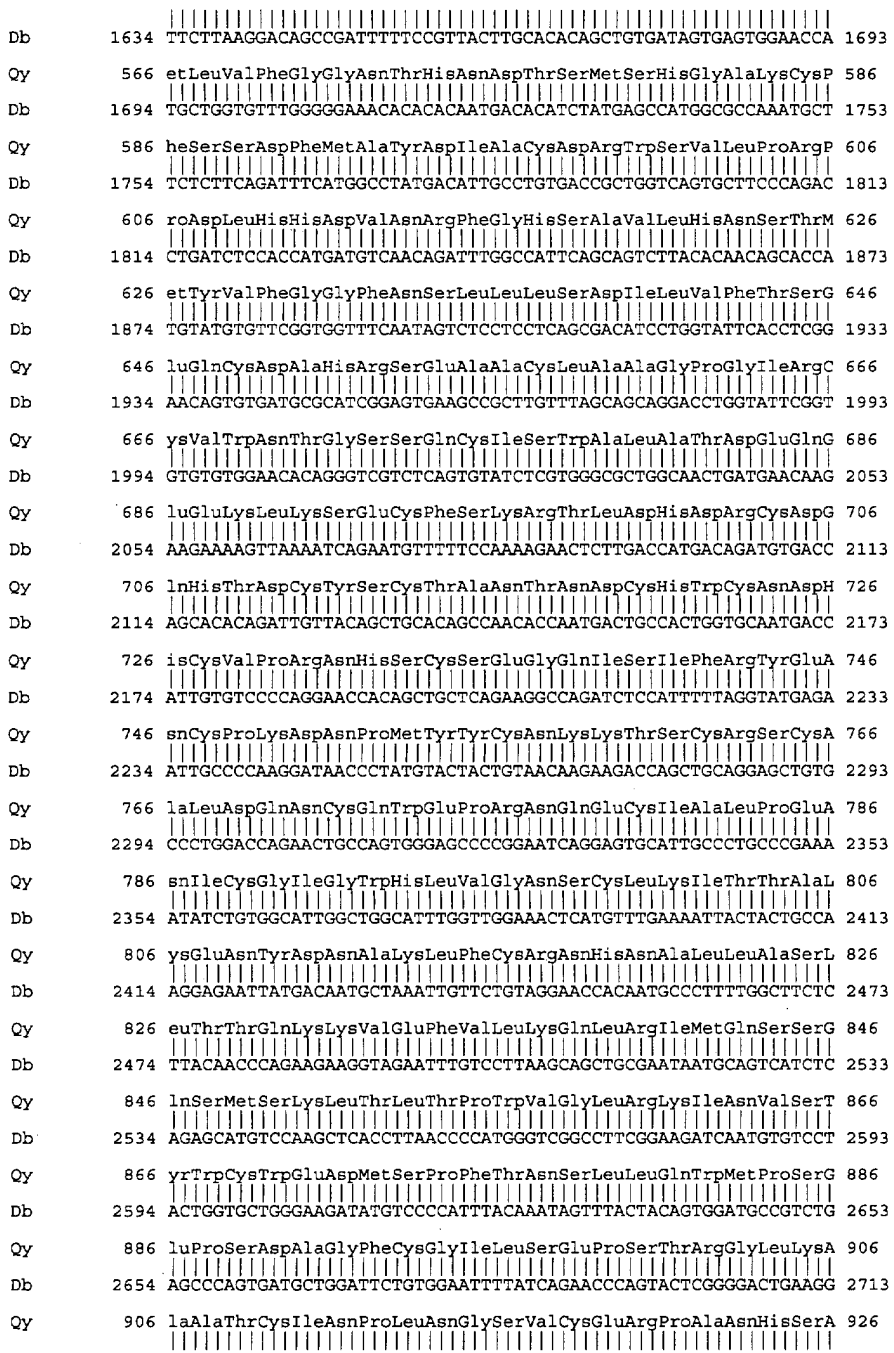
Qy       466 lyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsnT 486
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1394 GTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACA 1453

Qy       486 hrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisSerServ 506
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1454 CATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTG 1513

Qy       506 alTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaA 526
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1514 TTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCCA 1573

Qy       526 snLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetTrpThrI 546
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1574 ATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCA 1633

Qy       546 leLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerGlyThrM 566
```


Db 1634 TTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACCA 1693
Qy 566 etLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysP 586
Db 1694 TGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATGCT 1753
Qy 586 heSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuProArgP 606
Db 1754 TCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGAC 1813
Qy 606 roAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrM 626
Db 1814 CTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACAACAGCACCA 1873
Qy 626 etTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSerG 646
Db 1874 TGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCGG 1933
Qy 646 luGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgC 666
Db 1934 AACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGT 1993
Qy 666 ysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnG 686
Db 1994 GTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAAG 2053
Qy 686 luGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArgCysAspG 706
Db 2054 AAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACC 2113
Qy 706 lnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAspH 726
Db 2114 AGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGCAATGACC 2173
Qy 726 isCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGluA 746
Db 2174 ATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGA 2233
Qy 746 snCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArgSerCysA 766
Db 2234 ATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTG 2293
Qy 766 laLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuProGluA 786
Db 2294 CCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAA 2353
Qy 786 snIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThrThrAlaL 806
Db 2354 ATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCA 2413
Qy 806 ysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSerL 826
Db 2414 AGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTC 2473
Qy 826 euThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGlnSerSerG 846
Db 2474 TTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTC 2533
Qy 846 lnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerT 866
Db 2534 AGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCT 2593
Qy 866 yrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerG 886
Db 2594 ACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTG 2653
Qy 886 luProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysA 906
Db 2654 AGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGG 2713
Qy 906 laAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSerA 926
Db 2714 CTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTG 2773
Qy 926 laLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGlyS 946
Db 2774 CTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGGCA 2833
Qy 946 erSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrValA 966
Db 2834 GCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGTGG 2893
Qy 966 laSerProProPheGlyGlnCysMetGluTrpTyrThrMetSerThrCysProProGluA 986
Db 2894 CCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGAAA 2953
Qy 986 snCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCysT 1006
Db 2954 ATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTA 3013
Qy 1006 hrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyProValL 1026
Db 3014 CTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGA 3073
Qy 1026 ysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerSerM 1046
Db 3074 AGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGCA 3133
Qy 1046 etCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnCysA 1066
Db 3134 TGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCA 3193
Qy 1066 snGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrThrG 1086
Db 3194 ACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCACAG 3253
Qy 1086 lyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyGlyLysC 1106
Db 3254 GCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAAT 3313
Qy 1106 ysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysP 1126
Db 3314 GTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGCT 3373
Qy 1126 heCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGluAsnArgT 1146
Db 3374 TCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGAT 3433
Qy 1146 yrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrGlnPheT 1166
Db 3434 ACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTTCA 3493
Qy 1166 hrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIleAsnPheValAlaThrP 1186
Db 3494 CCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTC 3553
Qy 1186 roAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeuA 1206
Db 3554 CTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCTCA 3613
Qy 1206 snIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProValV 1226
Db 3614 ACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTG 3673
Qy 1226 alSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPheAspPheA 1246
Db 3674 TTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTC 3733
Qy 1246 rgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpProIleLysI 1266
Db 3734 GCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAAAA 3793
Qy 1266 leGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePheValThrP 1286
Db 3794 TTCAGATTGCCTTCTCTCAGCACAGCAATTTTATGGACCTGGTACAGTTCTTCGTGACTT 3853

```
Qy       1286 hePheSerCysPheLeuSerLeuLeuLeuValAlaAlaValValTrpLysIleLysGlnS 1306
Db       3854 TCTTCAGTTGTTTCCTCTCTTTGCTCCTGGTGGCTGCTGTGGTTTGGAAGATCAAACAAA 3913

Qy       1306 erCysTrpAlaSerArgArgArgGluGlnLeuLeuArgGluMetGlnGlnMetAlaSerA 1326
Db       3914 GTTGTTGGGCCTCCAGACGTAGAGAGCAACTTCTTCGAGAGATGCAACAGATGGCCAGCC 3973

Qy       1326 rgProPheAlaSerValAsnValAlaLeuGluThrAspGluGluProProAspLeuIleG 1346
Db       3974 GTCCCTTTGCCTCTGTAAATGTCGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTATTG 4033

Qy       1346 lyGlySerIleLysThrValProLysProIleAlaLeuGluProCysPheGlyAsnLysA 1366
Db       4034 GGGGGAGTATAAAGACTGTTCCCAAACCCATTGCACTGGAGCCGTGTTTTGGCAACAAAG 4093

Qy       1366 laAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyGlyIleProProProG 1386
Db       4094 CCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCCTGGGTGGCATCCCTCCTCCTG 4153

Qy       1386 lyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetProIleV 1406
Db       4154 GGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGATAG 4213

Qy       1406 alTyrLysGluLysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnProG 1426
Db       4214 TGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGCCTG 4273

Qy       1426 lyThrCysIle 1429
Db       4274 GGACCTGCATC 4284
```

RESULT 2
US-09-245-041-16
; Sequence 16, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 4072
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-245-041-16

Alignment Scores:

Pred. No.:	0	Length:	4072
Score:	979.00	Matches:	1179
Percent Similarity:	99.66%	Conservative:	0
Best Local Similarity:	99.66%	Mismatches:	2
Query Match:	68.51%	Indels:	4
DB:	3	Gaps:	0

US-09-787-097-12 (1-1429) x US-09-245-041-16 (1-4072)

```
Qy         87 AlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGluCysAspArgProCysVal 106
Db        256 GCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAAGGAATGTGACCGGCCGTGTGTC 315

Qy        107 AsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTrpValGly 126
Db        316 AACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGC 375

Qy        127 GluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPheValThr 146
Db        376 GAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACA 435

Qy        147 AspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeuIleGluGlyGlnPro 166
Db        436 GATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCA 495

Qy        167 AsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLeu 186
Db        496 AATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCATTTA 555

Qy        187 TyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaAlaPheSerGlyLeuIle 206
Db        556 TATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATT 615

Qy        207 ValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGlyTyrAla 226
Db        616 GTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCC 675

Qy        227 LeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThrTyrSer 246
Db        676 TTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTACAGT 735

Qy        247 PheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSerSer 266
Db        736 TTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGC 795

Qy        267 Glu-ThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIleProHi 286
Db        796 GA-TACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCA 854

Qy        286 sCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspValArgGl 306
Db        855 CTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGG 914

Qy        306 yCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAlaAsnGl 326
Db        915 ATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCA 974

Qy        326 nSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHisLysAl 346
Db        975 GTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGC 1034

Qy        346 aValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHisSerAspTy 366
Db       1035 TGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTA 1094

Qy        366 rAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSerVa 386
Db       1095 TAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGT 1154

Qy        386 lAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyrMe 406
Db       1155 GAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACAT 1214

Qy        406 tTyrGlyGlyLysIleAspPro-ThrGlyAsnValThrAsnGluLeuArgValPheHisI 426
Db       1215 GTATGGAGGAAAAATTGATTC-AACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACA 1273

Qy        426 leHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaValValG 446
Db       1274 TTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTG 1333

Qy        446 lyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValIlePheG 466
Db       1334 GGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTG 1393

Qy        466 lyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsnT 486
Db       1394 GTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACA 1453
```

```
Qy      486 hrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisSerServ 506
Db     1454 CATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTG 1513

Qy      506 alTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaA 526
Db     1514 TTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCCA 1573

Qy      526 snLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetTrpThrI 546
Db     1574 ATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCA 1633

Qy      546 leLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerGlyThrM 566
Db     1634 TTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACCA 1693

Qy      566 etLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysP 586
Db     1694 TGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATGCT 1753

Qy      586 heSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuProArgP 606
Db     1754 TCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGAC 1813

Qy      606 roAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrM 626
Db     1814 CTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACAACAGCACCA 1873

Qy      626 etTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSerG 646
Db     1874 TGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCGG 1933

Qy      646 luGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgC 666
Db     1934 AACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGT 1993

Qy      666 ysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnG 686
Db     1994 GTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAAG 2053

Qy      686 luGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArgCysAspG 706
Db     2054 AAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACC 2113

Qy      706 lnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAspH 726
Db     2114 AGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGCAATGACC 2173

Qy      726 isCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGluA 746
Db     2174 ATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGA 2233

Qy      746 snCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArgSerCysA 766
Db     2234 ATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTG 2293

Qy      766 laLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuProGluA 786
Db     2294 CCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCCGAAA 2353

Qy      786 snIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThrThrAlaL 806
Db     2354 ATATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAAATTACTACTGCCA 2413

Qy      806 ysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSerL 826
Db     2414 AGGAGAATTATGACAATGCTAAATTGTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTC 2473

Qy      826 euThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGlnSerSerG 846
Db     2474 TTACAACCCAGAAGAAGGTAGAATTTGTCCTTAAGCAGCTGCGAATAATGCAGTCATCTC 2533

Qy      846 lnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerT 866
Db     2534 AGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGCCTTCGGAAGATCAATGTGTCCT 2593

Qy      866 yrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerG 886
Db     2594 ACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTG 2653

Qy      886 luProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysA 906
Db     2654 AGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGG 2713

Qy      906 laAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSerA 926
Db     2714 CTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTG 2773

Qy      926 laLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGlyS 946
Db     2774 CTAAGCAGTGCCGGACACCATGTGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGGCA 2833

Qy      946 erSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrValA 966
Db     2834 GCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCCTATGTGG 2893

Qy      966 laSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSerThrCysProProGluA 986
Db     2894 CCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTGAAA 2953

Qy      986 snCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCysT 1006
Db     2954 ATTGTTCAGGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTA 3013

Qy     1006 hrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyProValL 1026
Db     3014 CTGATCCCAGCAATACTGGCAAAGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGA 3073

Qy     1026 ysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerSerM 1046
Db     3074 AGATGCCTTCGCAAGCCCCTACAGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGCA 3133

Qy     1046 etCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnCysA 1066
Db     3134 TGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCA 3193

Qy     1066 snGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrThrG 1086
Db     3194 ACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCACAG 3253

Qy     1086 lyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyGlyLysC 1106
Db     3254 GCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAAT 3313

Qy     1106 ysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysP 1126
Db     3314 GTCAGCCATGCAAGTGCAATGGGCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGCT 3373

Qy     1126 heCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGluAsnArgT 1146
Db     3374 TCTGCACCACCAAGGGCGTCAAGGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGAT 3433

Qy     1146 yrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrGlnPheT 1166
Db     3434 ACCAAGGAAACCCTCTCAGAGGAACATGTTATTATACTCTTCTTATTGACTATCAGTTCA 3493

Qy     1166 hrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIleAsnPheValAlaThrP 1186
Db     3494 CCTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTC 3553

Qy     1186 roAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeuA 1206
Db     3554 CTGACGAACAAAACAGGGATTTGGACATGTTCATCAATGCCTCCAAGAATTTCAACCTCA 3613

Qy     1206 snIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProValV 1226
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3614 ACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTG 3673

Qy       1226 alSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPheAspPheA 1246
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3674 TTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTC 3733

Qy       1246 rgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpProIleLysI 1266
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3734 GCAACCACCCAAATATCACTTTCTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAAAA 3793

Qy       1266 leGln 1267
                |||||
Db       3794 TTCAG 3798

RESULT 3
US-09-245-041-12
; Sequence 12, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT: Nagle, D.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
;  TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
;  FILE REFERENCE: 7853-136
;  CURRENT APPLICATION NUMBER: US/09/245,041
;  CURRENT FILING DATE: 1999-02-05
;  EARLIER APPLICATION NUMBER: 60/093,630
;  EARLIER FILING DATE: 1998-07-21
;  EARLIER APPLICATION NUMBER: 60/104,978
;  EARLIER FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 131
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 12
;   LENGTH: 6370
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-245-041-12

Alignment Scores:
Pred. No.:                0              Length:          6370
Score:                    503.00         Matches:         516
Percent Similarity:       99.81%         Conservative:    0
Best Local Similarity:    99.81%         Mismatches:      0
Query Match:              35.20%         Indels:          1
DB:                       3              Gaps:            0

US-09-787-097-12 (1-1429) x US-09-245-041-12 (1-6370)

Qy        914 AsnGlySerValCysGluArgProAlaAsnHisSerAlaLys-GlnCysArgThrProCy 933
                ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        506 AATGGTAGTGTCTGTGAAAGGCCTGCAAACCACAGTGCTAAGGCAGTGCCGGACACCATG 565

Qy        933 sAlaLeuArgThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSe 953
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 TGCCTTGAGGACAGCATGTGGAGATTGCACCAGCGGCAGCTCTGAGTGCATGTGGTGCAG 625

Qy        953 rAsnMetLysGlnCysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCy 973
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 CAACATGAAGCAGTGTGTGGACTCCAATGCCTATGTGGCCTCCTTCCCTTTTGGCCAGTG 685

Qy        973 sMetGluTrpTyrThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCy 993
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 TATGGAATGGTATACGATGAGCACCTGCCCCCCTGAAAATTGTTCAGGCTACTGTACCTG 745

Qy        993 sSerHisCysLeuGluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLy 1013
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 TAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGGTGTACTGATCCCAGCAATACTGGCAA 805

Qy       1013 sGlyLysCysIleGluGlySerTyrLysGlyProValLysMetProSerGlnAlaProTh 1033
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 AGGGAAATGCATAGAGGGTTCCTATAAAGGACCAGTGAAGATGCCTTCGCAAGCCCCTAC 865

Qy       1033 rGlyAsnPheTyrProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTy 1053
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 AGGAAATTTCTATCCACAGCCCCTGCTCAATTCCAGCATGTGTCTAGAGGACAGCAGATA 925

Qy       1053 rAsnTrpSerPheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAs 1073
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 CAACTGGTCTTTCATTCACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCATCAA 985

Qy       1073 nGlnSerIleCysGluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIl 1093
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 TCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCACAGGCAAGCACTGCGAGACCTGCAT 1045

Qy       1093 eSerGlyPheTyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGl 1113
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1046 ATCTGGCTTCTACGGTGATCCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCAATGG 1105

Qy       1113 yHisAlaSerLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLy 1133
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1106 GCACGCGTCTCTGTGCAACACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAA 1165

Qy       1133 sGlyAspGluCysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGl 1153
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1166 GGGGGACGAGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAGAGG 1225

Qy       1153 yThrCysTyrTyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAs 1173
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1226 AACATGTTATTATACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTATCCCAGGAAGA 1285

Qy       1173 pAspArgTyrTyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLe 1193
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1286 TGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCCTGACGAACAAAACAGGGATTT 1345

Qy       1193 uAspMetPheIleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPh 1213
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1346 GGACATGTTCATCAATGCCTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTT 1405

Qy       1213 eSerAlaGlyThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGl 1233
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1406 CTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAAGGA 1465

Qy       1233 uTyrLysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPh 1253
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1466 GTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCACCCAAATATCACTTT 1525

Qy       1253 ePheValTyrValSerAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHi 1273
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1526 CTTTGTTTATGTCAGTAATTTCACCTGGCCCATCAAAATTCAGATTGCCTTCTCTCAGCA 1585

Qy       1273 sSerAsnPheMetAspLeuValGlnPhePheValThrPhePheSerCysPheLeuSerLe 1293
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1586 CAGCAATTTTATGGACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCTTT 1645

Qy       1293 uLeuLeuValAlaAlaValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgAr 1313
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1646 GCTCCTGGTGGCTGCTGTGGTTTGGAAGATCAAACAAAGTTGTTGGGCCTCCAGACGTAG 1705

Qy       1313 gGluGlnLeuLeuArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnVa 1333
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1706 AGAGCAACTTCTTCGAGAGATGCAACAGATGGCCAGCCGTCCCTTTGCCTCTGTAAATGT 1765

Qy       1333 lAlaLeuGluThrAspGluGluProProAspLeuIleGlyGlySerIleLysThrValPr 1353
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1766 CGCCTTGGAAACAGATGAGGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCC 1825

Qy       1353 oLysProIleAlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheVa 1373
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1826 CAAACCCATTGCACTGGAGCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGT 1885

Qy       1373 lArgLeuProArgGlyLeuGlyGlyIleProProProGlyGlnSerGlyLeuAlaValAl 1393
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1886 GAGGCTCCCTCGAGGCCTGGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGC 1945

Qy       1393 aSerAlaLeuValAspIleSerGlnGlnMetProIleValTyrLysGluLysSerGlyAl 1413
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1946 CAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGC 2005

Qy       1413 aValArgAsnArgLysGlnGlnProProAlaGlnProGlyThrCysIle 1429
                ||||||||||||||||||||||||||||||||||||||||||||||||
Db       2006 CGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGCCTGGGACCTGCATC 2054

RESULT 4
US-09-245-041-18
; Sequence 18, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT: Nagle, D.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
;  TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
;  FILE REFERENCE: 7853-136
;  CURRENT APPLICATION NUMBER: US/09/245,041
;  CURRENT FILING DATE: 1999-02-05
;  EARLIER APPLICATION NUMBER: 60/093,630
;  EARLIER FILING DATE: 1998-07-21
;  EARLIER APPLICATION NUMBER: 60/104,978
;  EARLIER FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 131
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 18
;   LENGTH: 2625
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-245-041-18

Alignment Scores:
Pred. No.:                0              Length:          2625
Score:                    496.00         Matches:         696
Percent Similarity:       99.43%         Conservative:    0
Best Local Similarity:    99.43%         Mismatches:      2
Query Match:              34.71%         Indels:          4
DB:                       3              Gaps:            0

US-09-787-097-12 (1-1429) x US-09-245-041-18 (1-2625)

Qy         87 AlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGluCysAspArgProCysVal 106
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 GCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAAGGAATGTGACCGGCCGTGTGTC 315

Qy        107 AsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTrpValGly 126
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        316 AACGGCGGTCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCGCCGGCTGGGTGGGC 375

Qy        127 GluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPheValThr 146
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        376 GAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACA 435

Qy        147 AspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeuIleGluGlyGlnPro 166
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        436 GATGGACCTGGAAATTATAAATACAAAACGAAGTGCACGTGGCTCATTGAAGGACAGCCA 495

Qy        167 AsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLeu 186
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        496 AATAGAATAATGAGACTTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGACCATTTA 555

Qy        187 TyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaAlaPheSerGlyLeuIle 206
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 TATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGCCTCATT 615

Qy        207 ValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGlyTyrAla 226
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 GTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCC 675

Qy        227 LeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThrTyrSer 246
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 TTGCTGCATTTTTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTACAGT 735

Qy        247 PheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSerSer 266
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 TTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAGATCAGTAATAGCAGC 795

Qy        267 Glu-ThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIleProHi 286
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        796 GA-TACTGTTGAATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCA 854

Qy        286 sCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspValArgGl 306
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        855 CTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAAGTGATGTCAGAGG 914

Qy        306 yCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAlaAsnGl 326
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        915 ATGCTCCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTCAGTTCCTGTACCAGCTAACCA 974

Qy        326 nSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHisLysAl 346
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        975 GTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCCAGAGCATCTCATAAAGC 1034

Qy        346 aValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHisSerAspTy 366
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1035 TGTGGTCAATGGAAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTA 1094

Qy        366 rAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSerVa 386
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1095 TAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAACCGTTCTGT 1154

Qy        386 lAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyrMe 406
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 GAACAATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTACAT 1214

Qy        406 tTyrGlyGlyLysIleAspPro-ThrGlyAsnValThrAsnGluLeuArgValPheHisI 426
                ||||||||||||||||||    ||||||||||||||||||||||||||||||||||||||||
Db       1215 GTATGGAGGAAAAATTGATTC-AACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCACA 1273

Qy        426 leHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaValValG 446
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1274 TTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGGCAAAGGAGCAGTATGCAGTGGTTG 1333

Qy        446 lyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValIlePheG 466
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1334 GGCACTCTGCACACATTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTG 1393

Qy        466 lyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsnT 486
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1394 GTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACA 1453

Qy        486 hrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisSerServ 506
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1454 CATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTG 1513

Qy        506 alTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaA 526
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1514 TTTACGACCATAGGACCAGGGCCCTATACGTTCATGGTGGCTACAAGGCTTTCAGTGCCA 1573

Qy        526 snLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetTrpThrI 546
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1574 ATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCA 1633

Qy        546 leLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerGlyThrM 566
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1634 TTCTTAAGGACAGCCGATTTTTCCGTTACTTGCACACAGCTGTGATAGTGAGTGGAACCA 1693

Qy        566 etLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysP 586
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1694 TGCTGGTGTTTGGGGGAAACACACACAATGACACATCTATGAGCCATGGCGCCAAATGCT 1753

Qy        586 heSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuProArgP 606
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1754 TCTCTTCAGATTTCATGGCCTATGACATTGCCTGTGACCGCTGGTCAGTGCTTCCCAGAC 1813

Qy        606 roAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrM 626
```

```
Db       1814 CTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACAACAGCACCA 1873

Qy        626 etTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSerG 646

Db       1874 TGTATGTGTTCGGTGGTTTCAATAGTCTCCTCCTCAGCGACATCCTGGTATTCACCTCGG 1933

Qy        646 luGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgC 666

Db       1934 AACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTTAGCAGCAGGACCTGGTATTCGGT 1993

Qy        666 ysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnG 686

Db       1994 GTGTGTGGAACACAGGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAAG 2053

Qy        686 luGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArgCysAspG 706

Db       2054 AAGAAAAGTTAAAATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACC 2113

Qy        706 lnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAspH 726

Db       2114 AGCACACAGATTGTTACAGCTGCACAGCCAACACCAATGACTGCCACTGGTGCAATGACC 2173

Qy        726 isCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGluA 746

Db       2174 ATTGTGTCCCCAGGAACCACAGCTGCTCAGAAGGCCAGATCTCCATTTTTAGGTATGAGA 2233

Qy        746 snCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArgSerCysA 766

Db       2234 ATTGCCCCAAGGATAACCCTATGTACTACTGTAACAAGAAGACCAGCTGCAGGAGCTGTG 2293

Qy        766 laLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuPro 784

Db       2294 CCCTGGACCAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCC 2349

RESULT 5
US-09-245-041-1
; Sequence 1, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT: Nagle, D.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
;  TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
;  FILE REFERENCE: 7853-136
;  CURRENT APPLICATION NUMBER: US/09/245,041
;  CURRENT FILING DATE: 1999-02-05
;  EARLIER APPLICATION NUMBER: 60/093,630
;  EARLIER FILING DATE: 1998-07-21
;  EARLIER APPLICATION NUMBER: 60/104,978
;  EARLIER FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 131
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 8827
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-245-041-1

Alignment Scores:
Pred. No.:                3.2e-182        Length:          8827
Score:                    206.00          Matches:         576
Percent Similarity:       96.00%          Conservative:    0
Best Local Similarity:    96.00%          Mismatches:      12
Query Match:              14.42%          Indels:          24
DB:                       3               Gaps:            0

US-09-787-097-12 (1-1429) x US-09-245-041-1 (1-8827)

Qy        842 MetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLys 861

Db       2790 ATGCAATCATCTCAAAGTATGTCCAAGCTCACTCTGACTCCATGGGTTGGTCTTCGGAAG 2849

Qy        862 IleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGln 881

Db       2850 ATCAATGTGTCTTACTGGTGCTGGGAGGATATGTCTCCATTCACAAATAGTTTGCTGCAG 2909

Qy        882 TrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThr 901

Db       2910 TGGATGCCATCTGAGCCCAGTGATGCTGGCTTCTGTGGGATCTTGTCAGAGCCTAGTACT 2969

Qy        902 ArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro 921

Db       2970 CGGGGATTAAAGGCTGCAACCTGCATCAACCCTCTCAATGGCAGCGTCTGTGAAAGGCCT 3029

Qy        922 AlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAsp 941

Db       3030 GCAAACCACAGTGCCAAGCAGTGCCGGACACCATGTGCCCTGCGGACAGCGTGTGGCGA- 3088

Qy        942 -CysThrSerGly-SerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspS 961

Db       3089 GTGCACTAGCAG-CAGCTCGGAGTGCATGTGGTGCAGTAACATGAAGCAGTGTGTGGACT 3147

Qy        961 erAsnAlaTyrValAlaSerProProGlyGlnCysMetGluTrpTyrThrMetSerT 981

Db       3148 CCAATGCCTACGTGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCA 3207

Qy        981 hr-CysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPro 1000

Db       3208 G-CTGCCCACCTGAAAATTGCTCTGGCTACTGTACCTGCAGCCATTGCTTGGAGCAGCCA 3266

Qy       1001 GlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySer 1020

Db       3267 GGCTGTGGTTGGTGTACTGATCCTAGCAATACTGGGAAAGGAAAATGTATTGAGGGCAGC 3326

Qy       1021 TyrLysGlyProValLysMetProSerGlnAlaProThr--GlyAsnPhe-TyrProGln 1039

Db       3327 TATAAAGGACCTGTGAAGATGCCGTCACAGGCCTC-TG-CAGGAAATGT-GTATCCACAG 3383

Qy       1040 ProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHis 1059

Db       3384 CCCCTTCTGAACTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTCAC 3443

Qy       1060 CysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLys 1079

Db       3444 TGTCCAGCTTGCCAGTGCAACGGACACAGCAAATGCATCAACCAGAGTATCTGTGAGAAG 3503

Qy       1080 CysGluAsn-LeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAs 1099

Db       3504 TGTGAGGA-CCTGACCACGGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTATGGTGA 3562

Qy       1099 pProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAs 1119

Db       3563 CCCGACTAATGGAGGCAAATGTCAGCCATGCAAGTGCAATGGGCACGCATCACTGTGCAA 3622

Qy       1119 nThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLe 1139

Db       3623 CACCAACACCGGCAAGTGCTTCTGTACCACCAAAGGTGTCAAGGGGGACGAGTGCCAGCT 3682

Qy       1139 uCysGluValGluAsnArgTyrGlnGlyAsnProLeuArg-GlyThrCysTyrTyrThrL 1159

Db       3683 ATGTGAGGTAGAAAATCGATACCAAGGAAACCCTCTCAA-AGGAACATGCTACTATACCC 3741

Qy       1159 euLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrA 1179

Db       3742 TTCTCATTGACTATCAGTTCACCTTTAGCCTGTCCCAGGAAGACGACCGCTACTACACAG 3801

Qy       1179 laIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnA 1199

Db       3802 CCATCAACTTTGTGGCTACTCCTGATGAACAAAACAGGGATTTGGACATGTTCATCAATG 3861

Qy       1199 laSerLysAsnPheAsnLeuAsnIleThrTrpAlaAla-SerPheSer-AlaGlyThrGl 1218

Db       3862 CCTCCAAAAACTTCAACCTCAACATCACCTGGGCCAC-CAGCTTCCC-AGCCGGAACCCA 3919

Qy       1218 nAla-GlyGluGluMet-ProValValSerLysThrAsnIleLysGluTyrLysAspSer 1237
```

```
Db       3920 GAC-TGGAGAAGAGGT-GCCTGTTGTTTCAAAAACCAACATCAAGGAATACAAAGATAGC 3977

Qy       1238 PheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrVal 1257
Db       3978 TTCTCTAATGAGAAATTTGATTTTCGCAACCATCCAAACATCACTTTCTTTGTTTATGTC 4037

Qy       1258 SerAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMet 1277
Db       4038 AGTAATTTCACTTGGCCCATCAAAATTCAGATTGCCTTCTCCCAGCACAGCAACTTCATG 4097

Qy       1278 AspLeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuLeuValAla 1297
Db       4098 GACCTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTTCCTCTCGCTGCTTCTGGTGGCT 4157

Qy       1298 AlaValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeuLeu 1317
Db       4158 GCAGTGGTCTGGAAGATCAAGCAGAGCTGTTGGGCATCCAGGCGGAGAGAGCAACTTCTT 4217

Qy       1318 ArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThr 1337
Db       4218 CGGGAGATGCAACAGATGGCCAGCCGCCCCTTTGCTTCTGTAAACGTTGCCTTGGAAACA 4277

Qy       1338 AspGluGluProProAspLeuIleGlyGlySerIleLysThrValProLysProIleAla 1357
Db       4278 GATGAGGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCCAAACCCATTGCA 4337

Qy       1358 LeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProArg 1377
Db       4338 CTGGAGCCGTGTTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGGCTCCCTCGA 4397

Qy       1378 GlyLeuGlyGlyIleProProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuVal 1397
Db       4398 GGCCTGGGTGGCATCCCTCCTCCTGGGCAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTG 4457

Qy       1398 AspIleSerGlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAsnArg 1417
Db       4458 GACATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAACCGG 4517

Qy       1418 LysGlnGlnProProAlaGlnProGlyThrCysIle 1429
Db       4518 AAGCAGCAGCCCCCTGCACAGCCTGGGACCTGCATC 4553
```

RESULT 6
US-09-245-041-8
; Sequence 8, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 2419
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-8

Alignment Scores:

Pred. No.:	5.35e-78	Length:	2419
Score:	94.00	Matches:	94
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	6.58%	Indels:	0
DB:	3	Gaps:	0

US-09-787-097-12 (1-1429) x US-09-245-041-8 (1-2419)

```
Qy        513 AlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAsp 532
Db       1803 GCTCTGTACGTTCATGGTGGCTACAAGGCTTTCAGCGCCAACAAATACCGGCTTGCAGAT 1862

Qy        533 AspLeuTyrArgTyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPhe 552
Db       1863 GACCTCTACAGATACGATGTGGATACTCAGATGTGGACCATTCTTAAGGACAGCCGATTT 1922

Qy        553 PheArgTyrLeuHisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsn 572
Db       1923 TTCCGTTACTTGCATACAGCTGTGATAGTGAGTGGAACCATGCTGGTGTTTGGAGGGAAC 1982

Qy        573 ThrHisAsnAspThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAla 592
Db       1983 ACACACAATGACACTTCCATGAGCCACGGTGCCAAATGCTTCTCCTCGGACTTCATGGCT 2042

Qy        593 TyrAspIleAlaCysAspArgTrpSerValLeuProArgPro 606
Db       2043 TATGACATTGCTTGTGACCGATGGTCAGTGCTTCCCAGACCT 2084
```

RESULT 7
US-09-702-705-966/c
; Sequence 966, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-966

Alignment Scores:

Pred. No.:	9.18e-68	Length:	246
Score:	82.00	Matches:	82
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.74%	Indels:	0
DB:	4	Gaps:	0

US-09-787-097-12 (1-1429) x US-09-702-705-966 (1-246)

```
Qy        839 LeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGly 858
Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy        859 LeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSer 878
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy        879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGlu 898
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67
```

```
Qy        899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy        919 GluArg 920
              ||||||
Db          6 GAAAGG 1
```

RESULT 8
US-09-736-457-966/c
; Sequence 966, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-966

Alignment Scores:

Pred. No.:	9.18e-68	Length:	246
Score:	82.00	Matches:	82
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.74%	Indels:	0
DB:	4	Gaps:	0

US-09-787-097-12 (1-1429) x US-09-736-457-966 (1-246)

```
Qy        839 LeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGly 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy        859 LeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSer 878
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy        879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGlu 898
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67

Qy        899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy        919 GluArg 920
              ||||||
Db          6 GAAAGG 1
```

RESULT 9
US-09-614-124B-966/c
; Sequence 966, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-614-124B-966

Alignment Scores:

Pred. No.:	9.18e-68	Length:	246
Score:	82.00	Matches:	82
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.74%	Indels:	0
DB:	4	Gaps:	0

US-09-787-097-12 (1-1429) x US-09-614-124B-966 (1-246)

```
Qy        839 LeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGly 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy        859 LeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSer 878
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy        879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGlu 898
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67

Qy        899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy        919 GluArg 920
              ||||||
Db          6 GAAAGG 1
```

RESULT 10
US-09-671-325-966/c
; Sequence 966, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA

```
;   ORGANISM: Homo sapien
US-09-671-325-966

Alignment Scores:
Pred. No.:              9.18e-68        Length:         246
Score:                  82.00           Matches:        82
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            5.74%           Indels:         0
DB:                     4               Gaps:           0

US-09-787-097-12 (1-1429) x US-09-671-325-966 (1-246)

Qy        839 LeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGly 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC 187

Qy        859 LeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSer 878
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTCGGAAGATCAATGTGTCCTACTGGTGCTGGGAAGATATGTCCCCATTTACAAATAGT 127

Qy        879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGlu 898
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TTACTACAGTGGATGCCGTCTGAGCCCAGTGATGCTGGATTCTGTGGAATTTTATCAGAA 67

Qy        899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

Qy        919 GluArg 920
              ||||||
Db          6 GAAAGG 1

RESULT 11
US-09-245-041-10
; Sequence 10, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT: Nagle, D.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
;  TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
;  FILE REFERENCE: 7853-136
;  CURRENT APPLICATION NUMBER: US/09/245,041
;  CURRENT FILING DATE: 1999-02-05
;  EARLIER APPLICATION NUMBER: 60/093,630
;  EARLIER FILING DATE: 1998-07-21
;  EARLIER APPLICATION NUMBER: 60/104,978
;  EARLIER FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 131
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10
;   LENGTH: 1051
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-245-041-10

Alignment Scores:
Pred. No.:              1.61e-61        Length:         1051
Score:                  76.00           Matches:        111
Percent Similarity:     98.23%          Conservative:   0
Best Local Similarity:  98.23%          Mismatches:     1
Query Match:            5.32%           Indels:         2
DB:                     3               Gaps:           0

US-09-787-097-12 (1-1429) x US-09-245-041-10 (1-1051)

Qy         87 AlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGluCysAspArgProCysVal 106
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        463 GCGGCGGCGGTGTCCGGCTCGGCCGCAGCCGAGGCCAAGGAATGTGACCGGCCGTGTGTC 522

Qy        107 AsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCys-ProAlaGlyTrpValGl 126
              |||||||||||||||||||||||||||||||||||||||||| |||   ||||||||||
Db        523 AACGGCGGCCGCTGCAACCCTGGCACCGGCCAGTGCGTCTGCCCCACG-GGCTGGGTGGG 581

Qy        126 yGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPheValTh 146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        582 CGAGCAATGCCAGCACTGCGGGGGCCGCTTCAGACTAACTGGCTCTTCTGGATTTGTAAC 641

Qy        146 rAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeuIleGluGlyGlnPr 166
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        642 AGATGGACCTGGGAATTATAAATATAAGACGAAGTGCACATGGCTCATTGAAGGACAGCC 701

Qy        166 oAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLe 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        702 AAATAGAATAATGAGACTTCGCTTCAACCATTTTGCTACAGAATGTAGCTGGGACCATTT 761

Qy        186 uTyrValTyrAspGlyAspSerIleTyrAlaProLeu 198
              |||||||||||||||||||||||||||||||||||
Db        762 ATATGTTTATGATGGGGACTCAATCTACGCACCTCTG 798

RESULT 12
US-09-245-041-5
; Sequence 5, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT: Nagle, D.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
;  TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
;  FILE REFERENCE: 7853-136
;  CURRENT APPLICATION NUMBER: US/09/245,041
;  CURRENT FILING DATE: 1999-02-05
;  EARLIER APPLICATION NUMBER: 60/093,630
;  EARLIER FILING DATE: 1998-07-21
;  EARLIER APPLICATION NUMBER: 60/104,978
;  EARLIER FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 131
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5
;   LENGTH: 90050
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-245-041-5

Alignment Scores:
Pred. No.:              8.84e-57        Length:         90050
Score:                  73.00           Matches:        73
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            5.11%           Indels:         0
DB:                     3               Gaps:           0

US-09-787-097-12 (1-1429) x US-09-245-041-5 (1-90050)

Qy        849 SerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      28996 TCCAAGCTCACTCTGACTCCATGGGTTGGTCTTCGGAAGATCAATGTGTCTTACTGGTGC 29055

Qy        869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      29056 TGGGAGGATATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGT 29115

Qy        889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      29116 GATGCTGGCTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGATTAAAGGCTGCAACC 29175

Qy        909 CysIleAsnProLeuAsnGlySerValCysGluArgPro 921
              |||||||||||||||||||||||||||||||||||||
Db      29176 TGCATCAACCCTCTCAATGGCAGCGTCTGTGAAAGGCCT 29214

RESULT 13
US-09-245-041-6/c
; Sequence 6, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
```

```
;  APPLICANT: Moore, K.
;  APPLICANT: Nagle, D.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
;  TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
;  FILE REFERENCE: 7853-136
;  CURRENT APPLICATION NUMBER: US/09/245,041
;  CURRENT FILING DATE: 1999-02-05
;  EARLIER APPLICATION NUMBER: 60/093,630
;  EARLIER FILING DATE: 1998-07-21
;  EARLIER APPLICATION NUMBER: 60/104,978
;  EARLIER FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 131
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 6
;   LENGTH: 226
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-245-041-6

Alignment Scores:
Pred. No.:              8.49e-39        Length:         226
Score:                  51.00           Matches:        51
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            3.57%           Indels:         0
DB:                     3               Gaps:           0

US-09-787-097-12 (1-1429) x US-09-245-041-6 (1-226)

Qy        871 AspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSerAspAla 890
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        225 GATATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGTGATGCT 166

Qy        891 GlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThrCysIle 910
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        165 GGCTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGATTAAAGGCTGCAACCTGCATC 106

Qy        911 AsnProLeuAsnGlySerValCysGluArgPro 921
              |||||||||||||||||||||||||||||||
Db        105 AACCCTCTCAATGGCAGCGTCTGTGAAAGGCCT 73

RESULT 14
US-09-245-041-7/c
; Sequence 7, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT: Nagle, D.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
;  TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
;  FILE REFERENCE: 7853-136
;  CURRENT APPLICATION NUMBER: US/09/245,041
;  CURRENT FILING DATE: 1999-02-05
;  EARLIER APPLICATION NUMBER: 60/093,630
;  EARLIER FILING DATE: 1998-07-21
;  EARLIER APPLICATION NUMBER: 60/104,978
;  EARLIER FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 131
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7
;   LENGTH: 221
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-245-041-7

Alignment Scores:
Pred. No.:              3.41e-33        Length:         221
Score:                  45.00           Matches:        45
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            3.15%           Indels:         0
DB:                     3               Gaps:           0

US-09-787-097-12 (1-1429) x US-09-245-041-7 (1-221)

Qy        871 AspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSerAspAla 890
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 GATATGTCTCCATTCACAAATAGTTTGCTGCAGTGGATGCCATCTGAGCCCAGTGATGCT 161

Qy        891 GlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThrCysIle 910
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 GGCTTCTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGATTAAAGGCTGCAACCTGCATC 101

Qy        911 AsnProLeuAsnGly 915
              |||||||||||||||
Db        100 AACCCTCTCAATGGT 86

RESULT 15
US-09-245-041-4
; Sequence 4, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT: Nagle, D.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
;  TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
;  FILE REFERENCE: 7853-136
;  CURRENT APPLICATION NUMBER: US/09/245,041
;  CURRENT FILING DATE: 1999-02-05
;  EARLIER APPLICATION NUMBER: 60/093,630
;  EARLIER FILING DATE: 1998-07-21
;  EARLIER APPLICATION NUMBER: 60/104,978
;  EARLIER FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 131
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
;   LENGTH: 5973
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-245-041-4

Alignment Scores:
Pred. No.:              1.55e-20        Length:         5973
Score:                  33.00           Matches:        33
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            2.31%           Indels:         0
DB:                     3               Gaps:           0

US-09-787-097-12 (1-1429) x US-09-245-041-4 (1-5973)

Qy        166 ProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHis 185
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4328 CCAAATAGAATAATGAGACTTCGCTTCAACCATTTTGCTACAGAATGTAGCTGGGACCAT 4387

Qy        186 LeuTyrValTyrAspGlyAspSerIleTyrAlaProLeu 198
              |||||||||||||||||||||||||||||||||||||
Db       4388 TTATATGTTTATGATGGGGACTCAATCTACGCACCTCTG 4426

RESULT 16
US-09-245-041-3/c
; Sequence 3, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT: Nagle, D.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
;  TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
;  FILE REFERENCE: 7853-136
;  CURRENT APPLICATION NUMBER: US/09/245,041
;  CURRENT FILING DATE: 1999-02-05
;  EARLIER APPLICATION NUMBER: 60/093,630
;  EARLIER FILING DATE: 1998-07-21
;  EARLIER APPLICATION NUMBER: 60/104,978
;  EARLIER FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 131
```

```
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 17056
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-245-041-3

Alignment Scores:
Pred. No.:              2.84e-17        Length:         17056
Score:                  30.00           Matches:        30
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            2.10%           Indels:         0
DB:                     3               Gaps:           0

US-09-787-097-12 (1-1429) x US-09-245-041-3 (1-17056)

Qy        337 LeuLysLeuProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValVal 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2473 TTAAAGCTTCCCAGAGCCTCTCATAAAGCTGTGGTCAATGGAAATATAATGTGGGTTGTT 2414

Qy        357 GlyGlyTyrMetPheAsnHisSerAspTyr 366
              |||||||||||||||||||||||||||||||
Db       2413 GGCGGATATATGTTCAACCATTCAGATTAC 2384

RESULT 17
US-09-833-381-1916
; Sequence 1916, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
;  APPLICANT: Robison, Keith E.
;  TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
;  FILE REFERENCE: 5800-119
;  CURRENT APPLICATION NUMBER: US/09/833,381
;  CURRENT FILING DATE:  2001-04-11
;  PRIOR APPLICATION NUMBER: 09/516,448
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 2050
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1916
;   LENGTH: 415
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(415)
;   OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1916

Alignment Scores:
Pred. No.:              2.11e-11        Length:         415
Score:                  22.00           Matches:        22
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            1.54%           Indels:         0
DB:                     4               Gaps:           0

US-09-787-097-12 (1-1429) x US-09-833-381-1916 (1-415)

Qy        222 ThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPhe 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 ACATCTGGCTATGCACTGTTACATTTTTTTAGTGATGCTGCGTATAATCTAACTGGTTTC 306

Qy        242 AsnIle 243
              ||||||
Db        307 AACATT 312

RESULT 18
US-08-905-223-61
; Sequence 61, Application US/08905223
; Patent No. 6222029
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Jean-Baptiste D.
;    APPLICANT:  Duelert, Aymeric
;    APPLICANT:  Lacroix, Bruno
;    TITLE OF INVENTION:  5' ESTs FOR SECRETED PROTEINS
;    NUMBER OF SEQUENCES:  503
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Knobbe, Martens, Olson & Bear
;      STREET:  501 West Broadway
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-3505
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Win95
;      SOFTWARE:  Word
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/905,223
;      FILING DATE:
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Israelsen, Ned A.
;      REGISTRATION NUMBER:  29,655
;      REFERENCE/DOCKET NUMBER:
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 235-8550
;      TELEFAX:  (619) 235-0176
;  INFORMATION FOR SEQ ID NO:  61:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  429 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  DOUBLE
;      TOPOLOGY:  LINEAR
;    MOLECULE TYPE:  CDNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo Sapiens
;      TISSUE TYPE:  Brain
;    FEATURE:
;      NAME/KEY:  sig_peptide
;      LOCATION:  208..264
;      IDENTIFICATION METHOD:  Von Heijne matrix
;      OTHER INFORMATION:  score 7.6
;      OTHER INFORMATION:  seq LLNLLLLSLFAGL/DP
US-08-905-223-61

Alignment Scores:
Pred. No.:              0.00575         Length:         429
Score:                  13.00           Matches:        13
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            0.91%           Indels:         0
DB:                     3               Gaps:           0

US-09-787-097-12 (1-1429) x US-08-905-223-61 (1-429)

Qy         61 LeuLeuLeuLeuLeuLeuLeuProProProLeuLeu 73
              |||||||||||||||||||||||||||||||||||||
Db         54 CTGCTGTTATTGCTACTGCTGCTGCCGCCGCCTCTGCTT 92

RESULT 19
US-09-491-522-6
; Sequence 6, Application US/09491522
; Patent No. 6428998
;  GENERAL INFORMATION:
;    APPLICANT:  Colige, Alain
;    APPLICANT:  Lapiere, Charles M.
;    APPLICANT:  Prockop, Darwin J.
;    TITLE OF INVENTION:  RECOMBINANT N-PROTEINASE,
;    TITLE OF INVENTION:  AND THE PRODUCTION, METHODS AND USES THEREOF
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds, LLP
;      STREET:  1155 Avenue of the Americas
```

```
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/491,522
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/886,333
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Abrams, Samuel B
;      REGISTRATION NUMBER:  30,605
;      REFERENCE/DOCKET NUMBER:  8389-0060-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2023 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-491-522-6

Alignment Scores:
Pred. No.:                0.0271          Length:          2023
Score:                    13.00           Matches:         13
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              0.91%           Indels:          0
DB:                       4               Gaps:            0

US-09-787-097-12 (1-1429) x US-09-491-522-6 (1-2023)

Qy          61 LeuLeuLeuLeuLeuLeuLeuProProProLeuLeu 73
               |||||||||||||||||||||||||||||||||||||
Db         127 CTGCTGCTGCTGCTGCTGCTGCCGCCGCCGCTCCTG 165

RESULT 20
US-09-491-522-2
; Sequence 2, Application US/09491522
; Patent No. 6428998
;  GENERAL INFORMATION:
;    APPLICANT:  Colige, Alain
;    APPLICANT:  Lapiere, Charles M.
;    APPLICANT:  Prockop, Darwin J.
;    TITLE OF INVENTION:  RECOMBINANT N-PROTEINASE,
;    TITLE OF INVENTION:  AND THE PRODUCTION, METHODS AND USES THEREOF
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds, LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/491,522
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/886,333
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Abrams, Samuel B
;      REGISTRATION NUMBER:  30,605
;      REFERENCE/DOCKET NUMBER:  8389-0060-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2450 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-491-522-2

Alignment Scores:
Pred. No.:                0.0328          Length:          2450
Score:                    13.00           Matches:         13
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              0.91%           Indels:          0
DB:                       4               Gaps:            0

US-09-787-097-12 (1-1429) x US-09-491-522-2 (1-2450)

Qy          61 LeuLeuLeuLeuLeuLeuLeuProProProLeuLeu 73
               |||||||||||||||||||||||||||||||||||||
Db         127 CTGCTGCTGCTGCTGCTGCTGCCGCCGCCGCTCCTG 165

RESULT 21
US-09-491-522-1
; Sequence 1, Application US/09491522
; Patent No. 6428998
;  GENERAL INFORMATION:
;    APPLICANT:  Colige, Alain
;    APPLICANT:  Lapiere, Charles M.
;    APPLICANT:  Prockop, Darwin J.
;    TITLE OF INVENTION:  RECOMBINANT N-PROTEINASE,
;    TITLE OF INVENTION:  AND THE PRODUCTION, METHODS AND USES THEREOF
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds, LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/491,522
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/886,333
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Abrams, Samuel B
;      REGISTRATION NUMBER:  30,605
;      REFERENCE/DOCKET NUMBER:  8389-0060-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
```

```
;      LENGTH:  6692 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-491-522-1

Alignment Scores:
Pred. No.:              0.0896          Length:         6692
Score:                  13.00           Matches:        13
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            0.91%           Indels:         0
DB:                     4               Gaps:           0

US-09-787-097-12 (1-1429) x US-09-491-522-1 (1-6692)

Qy         61 LeuLeuLeuLeuLeuLeuLeuProProProLeuLeu 73
              |||||||||||||||||||||||||||||||||||||
Db        127 CTGCTGCTGCTGCTGCTGCTGCCGCCGCCGCTCCTG 165

RESULT 22
US-08-927-219-54
; Sequence 54, Application US/08927219
; Patent No. 6187533
;  GENERAL INFORMATION:
;    APPLICANT:  Bell, Graeme I.
;    APPLICANT:  Yamagata, Kazuya
;    APPLICANT:  Oda, Naohisha
;    APPLICANT:  Kaisaki, Pamela J.
;    APPLICANT:  Furuta, Hiroto
;    APPLICANT:  Horikawa, Yukio
;    APPLICANT:  Menzel, Stephen
;    TITLE OF INVENTION:  MUTATIONS IN THE DIABETES SUSCEPTIBILITY
;    TITLE OF INVENTION:  GENES HEPATOCYTE NUCLEAR FACTOR (HNF) 1 ALPHA, HNF-1BETA
;    TITLE OF INVENTION:  AND HNF-4ALPHA
;    NUMBER OF SEQUENCES:  147
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/927,219
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/029,679
;      FILING DATE:  30-OCT-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/028,056
;      FILING DATE:  02-OCT-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/025,719
;      FILING DATE:  10-SEP-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wilson, Mark B.
;      REGISTRATION NUMBER:  37,259
;      REFERENCE/DOCKET NUMBER:  ARCD:272
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/418-3000
;      TELEFAX:  512/474-7577
;  INFORMATION FOR SEQ ID NO:  54:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1103 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(289..429, 433..477, 481..492, 496..603, 607
;      LOCATION:  ..630, 634..750, 754..810, 814..843, 847..1023,
;      LOCATION:  1027..1071, 1075..1103)
US-08-927-219-54

Alignment Scores:
Pred. No.:              0.127           Length:         1103
Score:                  12.00           Matches:        12
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            0.84%           Indels:         0
DB:                     3               Gaps:           0

US-09-787-097-12 (1-1429) x US-08-927-219-54 (1-1103)

Qy         62 LeuLeuLeuLeuLeuLeuProProProLeuLeu 73
              |||||||||||||||||||||||||||||||||||||
Db        878 TTGCTCCTCCTTCTGCTGCTGCCCCCACCTCTGCTG 913

RESULT 23
US-09-028-327-2
; Sequence 2, Application US/09028327
; Patent No. 6130064
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Y. Tom
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Guegler, Karl J.
;    TITLE OF INVENTION:  HUMAN SMN-LIKE PROTEIN
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/028,327
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0482 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2426 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  BRSTNOT24
;      CLONE:  3769729
US-09-028-327-2

Alignment Scores:
Pred. No.:              0.28            Length:         2426
Score:                  12.00           Matches:        12
```

```
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            0.84%          Indels:        0
DB:                     3              Gaps:          0

US-09-787-097-12 (1-1429) x US-09-028-327-2 (1-2426)

Qy         61 LeuLeuLeuLeuLeuLeuLeuProProProLeu 72
              ||||||||||||||||||||||||||||||||||
Db        507 CTGCTACTGCTACTGCTGCTGCCACCGCCACTA 542

RESULT 24
US-09-571-078A-2
; Sequence 2, Application US/09571078A
; Patent No. 6620783
;   GENERAL INFORMATION:
;        APPLICANT: Tang, Y. Tom
;                   Corley, Neil C.
;                   Guegler, Karl J.
;        TITLE OF INVENTION: HUMAN SMN-LIKE PROTEIN
;        NUMBER OF SEQUENCES: 4
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Incyte Pharmaceuticals, Inc.
;             STREET: 3174 Porter Dr.
;             CITY: Palo Alto
;             STATE: CA
;             COUNTRY: USA
;             ZIP: 94304
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: DOS
;             SOFTWARE: FastSEQ for Windows Version 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/571,078A
;             FILING DATE: 15-May-2000
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: <Unknown>
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Billings, Lucy J.
;             REGISTRATION NUMBER: 36,749
;             REFERENCE/DOCKET NUMBER: PF-0482 US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 650-855-0555
;             TELEFAX: 650-845-4166
;             TELEX: <Unknown>
;   INFORMATION FOR SEQ ID NO: 2:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2426 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        IMMEDIATE SOURCE:
;             LIBRARY: BRSTNOT24
;             CLONE: 3769729
;        SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-571-078A-2

Alignment Scores:
Pred. No.:              0.28           Length:        2426
Score:                  12.00          Matches:       12
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            0.84%          Indels:        0
DB:                     4              Gaps:          0

US-09-787-097-12 (1-1429) x US-09-571-078A-2 (1-2426)

Qy         61 LeuLeuLeuLeuLeuLeuLeuProProProLeu 72
              ||||||||||||||||||||||||||||||||||
Db        507 CTGCTACTGCTACTGCTGCTGCCACCGCCACTA 542

RESULT 25
US-09-336-536-57
; Sequence 57, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
;  APPLICANT: Leiby, K.
;  APPLICANT: McKay, C.
;  APPLICANT: Bossone, S.
;  TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
;  FILE REFERENCE: 7853-144
;  CURRENT APPLICATION NUMBER: US/09/336,536
;  CURRENT FILING DATE: 1999-06-18
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 57
;   LENGTH: 639
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-336-536-57

Alignment Scores:
Pred. No.:              0.636          Length:        639
Score:                  11.00          Matches:       11
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            0.77%          Indels:        0
DB:                     4              Gaps:          0

US-09-787-097-12 (1-1429) x US-09-336-536-57 (1-639)

Qy         61 LeuLeuLeuLeuLeuLeuLeuProProPro 71
              |||||||||||||||||||||||||||||||
Db         31 CTATTGCTGCTGCTGTTGCTGCTGCCGCCGCCT 63

RESULT 26
US-09-620-312D-420/c
; Sequence 420, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  John Tillinghast
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. 6569662el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2B
;  CURRENT APPLICATION NUMBER: US/09/620,312D
;  CURRENT FILING DATE:  2000-07-19
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1105
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 420
;   LENGTH: 1432
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
```

; NAME/KEY: CDS
; LOCATION: (59)..(1030)
US-09-620-312D-420

Alignment Scores:

Pred. No.:	1.43	Length:	1432
Score:	11.00	Matches:	11
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.77%	Indels:	0
DB:	4	Gaps:	0

US-09-787-097-12 (1-1429) x US-09-620-312D-420 (1-1432)

```
Qy         64 LeuLeuLeuLeuLeuProProProLeuLeuLeu 74
              ||||||||||||||||||||||||||||||||||
Db        628 CTCCTCCTGCTGCTGCCTCCGCCGCTCCTCCTC 596
```

RESULT 27
US-09-148-545-11
; Sequence 11, Application US/09148545
; Patent No. 6590075
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: PZ001P1
; CURRENT APPLICATION NUMBER: US/09/148,545
; CURRENT FILING DATE: 1998-09-04
; EARLIER APPLICATION NUMBER: PCT/US98/04482
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,161
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,894
; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,911
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,636
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,874
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,910
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,864
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,631
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,845
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,892
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,595
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/057,761
; EARLIER FILING DATE: 05-Sep-1997
; EARLIER APPLICATION NUMBER: 60/047,599
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,588
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,586
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,590
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 280
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 1739

Alignment Scores:

Pred. No.:	1.73	Length:	1739
Score:	11.00	Matches:	11
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.77%	Indels:	0
DB:	4	Gaps:	0

US-09-787-097-12 (1-1429) x US-09-148-545-11 (1-1739)

```
Qy         61 LeuLeuLeuLeuLeuLeuLeuProProPro 71
              ||||||||||||||||||||||||||||||||
Db         93 CTGCTGTTGCTGTTGCTGCTGCCGCCGCCG 125
```

RESULT 28
US-09-336-536-56
; Sequence 56, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
; APPLICANT: Leiby, K.
; APPLICANT: McKay, C.
; APPLICANT: Bossone, S.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-144
; CURRENT APPLICATION NUMBER: US/09/336,536
; CURRENT FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 56
; LENGTH: 1858
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-336-536-56

Alignment Scores:

Pred. No.:	1.85	Length:	1858
Score:	11.00	Matches:	11
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.77%	Indels:	0
DB:	4	Gaps:	0

US-09-787-097-12 (1-1429) x US-09-336-536-56 (1-1858)

```
Qy         61 LeuLeuLeuLeuLeuLeuLeuProProPro 71
              ||||||||||||||||||||||||||||||||
Db        120 CTATTGCTGCTGCTGTTGCTGCTGCCGCCGCCT 152
```

RESULT 29
US-09-919-060-15
; Sequence 15, Application US/09919060
; Patent No. 6638744
; GENERAL INFORMATION:
; APPLICANT: Wisnewski, Nancy
; APPLICANT: Brandt, Kevin S.
; TITLE OF INVENTION: CANINE COX-1 AND COX-2 NUCLEIC ACID MOLECULES, PROTEINS AND USES
; FILE REFERENCE: AD-1
; CURRENT APPLICATION NUMBER: US/09/919,060
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/224,486
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 1899
; TYPE: DNA
; ORGANISM: Canis familiaris
US-09-919-060-15

Alignment Scores:

Pred. No.:	1.89	Length:	1899

```
Score:                  11.00          Matches:       11
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            0.77%          Indels:        0
DB:                     4              Gaps:          0

US-09-787-097-12 (1-1429) x US-09-919-060-15 (1-1899)

Qy         61 LeuLeuLeuLeuLeuLeuLeuProProPro 71
              |||||||||||||||||||||||||||||||
Db        127 CTCCTGCTGCTGCTGCTGCTGCCGCCGCCC 159

RESULT 30
US-09-919-060-16/c
; Sequence 16, Application US/09919060
; Patent No. 6638744
; GENERAL INFORMATION:
;  APPLICANT: Wisnewski, Nancy
;  APPLICANT:  Brandt, Kevin S.
;  TITLE OF INVENTION: CANINE COX-1 AND COX-2 NUCLEIC ACID MOLECULES, PROTEINS AND USES
;  FILE REFERENCE: AD-1
;  CURRENT APPLICATION NUMBER: US/09/919,060
;  CURRENT FILING DATE:  2001-07-31
;  PRIOR APPLICATION NUMBER: 60/224,486
;  PRIOR FILING DATE: 2000-08-11
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
;   LENGTH: 1899
;   TYPE: DNA
;   ORGANISM: Canis familiaris
US-09-919-060-16

Alignment Scores:
Pred. No.:              1.89           Length:        1899
Score:                  11.00          Matches:       11
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            0.77%          Indels:        0
DB:                     4              Gaps:          0

US-09-787-097-12 (1-1429) x US-09-919-060-16 (1-1899)

Qy         61 LeuLeuLeuLeuLeuLeuLeuProProPro 71
              |||||||||||||||||||||||||||||||
Db       1773 CTCCTGCTGCTGCTGCTGCTGCCGCCGCCC 1741

RESULT 31
US-09-907-794A-194
; Sequence 194, Application US/09907794A
; Patent No. 6635468
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/907,794A
;  CURRENT FILING DATE:  2001-07-17
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 194
;   LENGTH: 2362
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-907-794A-194

Alignment Scores:
Pred. No.:              2.35           Length:        2362
Score:                  11.00          Matches:       11
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            0.77%          Indels:        0
DB:                     4              Gaps:          0

US-09-787-097-12 (1-1429) x US-09-907-794A-194 (1-2362)

Qy         61 LeuLeuLeuLeuLeuLeuLeuProProPro 71
              |||||||||||||||||||||||||||||||
Db        106 CTGCTGTTGCTGTTGCTGCTGCCGCCGCCG 138

RESULT 32
US-09-905-125A-194
; Sequence 194, Application US/09905125A
; Patent No. 6664376
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
```

; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,125A
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 194
; LENGTH: 2362
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-905-125A-194

Alignment Scores:

Pred. No.:	2.35	Length:	2362
Score:	11.00	Matches:	11
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.77%	Indels:	0

DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-905-125A-194 (1-2362)

```
Qy         61 LeuLeuLeuLeuLeuLeuLeuProProPro 71
              |||||||||||||||||||||||||||||||
Db        106 CTGCTGTTGCTGTTGCTGCTGCTGCCGCCGCCG 138
```

RESULT 33
US-09-902-775A-194
; Sequence 194, Application US/09902775A
; Patent No. 6686451
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/902,775A
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20

```
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 194
;   LENGTH: 2362
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-902-775A-194

Alignment Scores:
Pred. No.:              2.35          Length:         2362
Score:                  11.00         Matches:        11
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            0.77%         Indels:         0
DB:                     4             Gaps:           0

US-09-787-097-12 (1-1429) x US-09-902-775A-194 (1-2362)

Qy          61 LeuLeuLeuLeuLeuLeuLeuProProPro 71
               |||||||||||||||||||||||||||||||
Db         106 CTGCTGTTGCTGTTGCTGCTGCTGCCGCCGCCG 138

RESULT 34
US-08-405-254-9/c
; Sequence 9, Application US/08405254
; Patent No. 5773288
;  GENERAL INFORMATION:
;    APPLICANT:  BRIGGS, STEVEN P.
;    APPLICANT:  BENSON, ROBERT J.
;    TITLE OF INVENTION:  PLANT GENES AFFECTING GIBBERELLIC ACID
;    TITLE OF INVENTION:  BIOSYNTHESIS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FOLEY & LARDNER
;      STREET:  3000 K STREET, NW, SUITE 500
;      CITY:  WASHINGTON
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20007
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/405,254
;      FILING DATE:  16-MAR-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/261,465
;      FILING DATE:  17-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, STEPHEN A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  33229/299/PIHI
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2520 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-405-254-9

Alignment Scores:
Pred. No.:              2.51          Length:         2520
Score:                  11.00         Matches:        11
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            0.77%         Indels:         0
DB:                     1             Gaps:           0

US-09-787-097-12 (1-1429) x US-08-405-254-9 (1-2520)

Qy          58 ArgProArgLeuLeuLeuLeuLeuLeuLeu 68
               |||||||||||||||||||||||||||||||
Db        2140 CGCCCGCGCCTCCTCCTCCTCCTGCTGCTGCTG 2108

RESULT 35
US-09-919-060-12
; Sequence 12, Application US/09919060
; Patent No. 6638744
; GENERAL INFORMATION:
;  APPLICANT: Wisnewski, Nancy
;  APPLICANT:  Brandt, Kevin S.
;  TITLE OF INVENTION: CANINE COX-1 AND COX-2 NUCLEIC ACID MOLECULES, PROTEINS AND USES
;  FILE REFERENCE: AD-1
;  CURRENT APPLICATION NUMBER: US/09/919,060
;  CURRENT FILING DATE:  2001-07-31
;  PRIOR APPLICATION NUMBER: 60/224,486
;  PRIOR FILING DATE: 2000-08-11
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
;   LENGTH: 2693
;   TYPE: DNA
;   ORGANISM: Canis familiaris
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (31)..(1929)
;   OTHER INFORMATION:
US-09-919-060-12

Alignment Scores:
Pred. No.:              2.68          Length:         2693
Score:                  11.00         Matches:        11
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            0.77%         Indels:         0
DB:                     4             Gaps:           0

US-09-787-097-12 (1-1429) x US-09-919-060-12 (1-2693)

Qy          61 LeuLeuLeuLeuLeuLeuLeuProProPro 71
               |||||||||||||||||||||||||||||||
Db        157 CTCCTGCTGCTGCTGCTGCTGCCGCCGCCC 189

RESULT 36
US-09-919-060-14/c
; Sequence 14, Application US/09919060
; Patent No. 6638744
; GENERAL INFORMATION:
;  APPLICANT: Wisnewski, Nancy
;  APPLICANT:  Brandt, Kevin S.
;  TITLE OF INVENTION: CANINE COX-1 AND COX-2 NUCLEIC ACID MOLECULES, PROTEINS AND USES
;  FILE REFERENCE: AD-1
;  CURRENT APPLICATION NUMBER: US/09/919,060
;  CURRENT FILING DATE:  2001-07-31
;  PRIOR APPLICATION NUMBER: 60/224,486
;  PRIOR FILING DATE: 2000-08-11
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
;   LENGTH: 2693
;   TYPE: DNA
;   ORGANISM: Canis familiaris
US-09-919-060-14

Alignment Scores:
Pred. No.:              2.68          Length:         2693
```

```
Score:                  11.00          Matches:       11
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            0.77%          Indels:        0
DB:                     4              Gaps:          0

US-09-787-097-12 (1-1429) x US-09-919-060-14 (1-2693)

Qy         61 LeuLeuLeuLeuLeuLeuLeuProProPro 71
              |||||||||||||||||||||||||||||||
Db       2537 CTCCTGCTGCTGCTGCTGCTGCCGCCGCCC 2505

RESULT 37
US-08-261-465-3/c
; Sequence 3, Application US/08261465
; Patent No. 5612191
;  GENERAL INFORMATION:
;    APPLICANT:  BRIGGS, Steven P.
;    APPLICANT:  BENSEN, Robert J.
;    TITLE OF INVENTION:  PLANT GENES AFFECTING GIBBERELLIC ACID
;    TITLE OF INVENTION:  BIOSYNTHESIS
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/261,465
;      FILING DATE:  17-JUN-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  33229/213/PIHI
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2784 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-261-465-3

Alignment Scores:
Pred. No.:              2.77           Length:        2784
Score:                  11.00          Matches:       11
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            0.77%          Indels:        0
DB:                     1              Gaps:          0

US-09-787-097-12 (1-1429) x US-08-261-465-3 (1-2784)

Qy         58 ArgProArgLeuLeuLeuLeuLeuLeuLeu 68
              |||||||||||||||||||||||||||||||
Db        171 CGCCCGCGCCTCCTCCTCCTCCTGCTGCTT 139

RESULT 38
US-08-405-254-7/c
; Sequence 7, Application US/08405254
; Patent No. 5773288
;  GENERAL INFORMATION:
;    APPLICANT:  BRIGGS, STEVEN P.
;    APPLICANT:  BENSON, ROBERT J.
;    TITLE OF INVENTION:  PLANT GENES AFFECTING GIBBERELLIC ACID
;    TITLE OF INVENTION:  BIOSYNTHESIS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FOLEY & LARDNER
;      STREET:  3000 K STREET, NW, SUITE 500
;      CITY:  WASHINGTON
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20007
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/405,254
;      FILING DATE:  16-MAR-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/261,465
;      FILING DATE:  17-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, STEPHEN A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  33229/299/PIHI
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2784 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-405-254-7

Alignment Scores:
Pred. No.:              2.77           Length:        2784
Score:                  11.00          Matches:       11
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            0.77%          Indels:        0
DB:                     1              Gaps:          0

US-09-787-097-12 (1-1429) x US-08-405-254-7 (1-2784)

Qy         58 ArgProArgLeuLeuLeuLeuLeuLeuLeu 68
              |||||||||||||||||||||||||||||||
Db        171 CGCCCGCGCCTCCTCCTCCTCCTGCTGCTT 139

RESULT 39
US-09-023-655-1072/c
; Sequence 1072, Application US/09023655
; Patent No. 6607879
;  GENERAL INFORMATION:
;    APPLICANT:  Cocks, Benjamin G.
;    APPLICANT:  Susan G. Stuart
;    APPLICANT:  Jeffrey J. Seilhamer
;    TITLE OF INVENTION:  COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
;    TITLE OF INVENTION:  EXPRESSION
;    NUMBER OF SEQUENCES:  1508
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/023,655
;       FILING DATE:  HEREWITH
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Zeller, Karen J.
;       REGISTRATION NUMBER:  37,071
;       REFERENCE/DOCKET NUMBER:  PA-0001 US
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (650) 855-0555
;       TELEFAX:  (650) 845-4166
;   INFORMATION FOR SEQ ID NO:  1072:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  3110 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     IMMEDIATE SOURCE:
;       LIBRARY:  GENBANK
;       CLONE:  g182262
US-09-023-655-1072

Alignment Scores:
Pred. No.:              3.1          Length:         3110
Score:                  11.00        Matches:        11
Percent Similarity:     100.00%      Conservative:   0
Best Local Similarity:  100.00%      Mismatches:     0
Query Match:            0.77%        Indels:         0
DB:                     4            Gaps:           0

US-09-787-097-12 (1-1429) x US-09-023-655-1072 (1-3110)

Qy          61 LeuLeuLeuLeuLeuLeuLeuProProPro 71
               |||||||||||||||||||||||||||||||
Db         505 CTGCTGCTGCTGCTGCTGTTGCTGCCGCCCCCG 473

RESULT 40
US-08-224-482-3/c
; Sequence 3, Application US/08224482
; Patent No. 5837692
;  GENERAL INFORMATION:
;    APPLICANT:  Mercola, Dan
;    APPLICANT:  Adamson, Eileen D.
;    TITLE OF INVENTION:  Inhibition of the Mitogenic Activity of
;    TITLE OF INVENTION:  PDGF by Mammalian EGR
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/224,482
;      FILING DATE:  07-APR-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-ME 9913
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3132 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-224-482-3

Alignment Scores:
Pred. No.:              3.12         Length:         3132
Score:                  11.00        Matches:        11
Percent Similarity:     100.00%      Conservative:   0
Best Local Similarity:  100.00%      Mismatches:     0
Query Match:            0.77%        Indels:         0
DB:                     2            Gaps:           0

US-09-787-097-12 (1-1429) x US-08-224-482-3 (1-3132)

Qy          61 LeuLeuLeuLeuLeuLeuLeuProProPro 71
               |||||||||||||||||||||||||||||||
Db         521 CTGCTGCTGCTGCTGCTGTTGCTGCCGCCCCCG 489

RESULT 41
US-09-205-921-1/c
; Sequence 1, Application US/09205921A
; Patent No. 6008048
; GENERAL INFORMATION:
;  APPLICANT: Brett P. Monia
;  APPLICANT: ex M. Cowsert
;  TITLE OF INVENTION: ANTISENSE MODULATION OF EGR-1 EXPRESSION
;  FILE REFERENCE: RTS-0028
;  CURRENT APPLICATION NUMBER: US/09/205,921A
;  CURRENT FILING DATE: 1998-12-04
;  NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 1
;   LENGTH: 3132
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (271)..(1902)
US-09-205-921-1

Alignment Scores:
Pred. No.:              3.12         Length:         3132
Score:                  11.00        Matches:        11
Percent Similarity:     100.00%      Conservative:   0
Best Local Similarity:  100.00%      Mismatches:     0
Query Match:            0.77%        Indels:         0
DB:                     3            Gaps:           0

US-09-787-097-12 (1-1429) x US-09-205-921-1 (1-3132)

Qy          61 LeuLeuLeuLeuLeuLeuLeuProProPro 71
               |||||||||||||||||||||||||||||||
Db         521 CTGCTGCTGCTGCTGCTGTTGCTGCCGCCCCCG 489

RESULT 42
US-09-300-958A-32/c
; Sequence 32, Application US/09300958A
; Patent No. 6495319
; GENERAL INFORMATION:
;  APPLICANT:  McClelland, Michael
;  APPLICANT:  Welsh, John
;  APPLICANT:  Trenkle, Thomas
;  TITLE OF INVENTION:  Reduced Complexity Nucleic Acid Targets and Methods of
;  TITLE OF INVENTION:  Using Same
;  FILE REFERENCE:  P-PH 3457
;  CURRENT APPLICATION NUMBER: US/09/300,958A
;  CURRENT FILING DATE:  1999-04-27
;  PRIOR APPLICATION NUMBER:  60/083,331
```

```
;  PRIOR FILING DATE:  1998-04-27
;  PRIOR APPLICATION NUMBER:  60/098,070
;  PRIOR FILING DATE:  1998-08-27
;  PRIOR APPLICATION NUMBER:  60/118,624
;  PRIOR FILING DATE:  1999-02-04
;  NUMBER OF SEQ ID NOS:  85
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 32
;   LENGTH: 3132
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-300-958A-32

Alignment Scores:
Pred. No.:              3.12          Length:          3132
Score:                  11.00         Matches:         11
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            0.77%         Indels:          0
DB:                     4             Gaps:            0

US-09-787-097-12 (1-1429) x US-09-300-958A-32 (1-3132)

Qy          61 LeuLeuLeuLeuLeuLeuLeuProProPro 71
               |||||||||||||||||||||||||||||||
Db         521 CTGCTGCTGCTGCTGCTGTTGCTGCCGCCCCCG 489

RESULT 43
US-09-702-705-320/c
; Sequence 320, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Fan, Liqun
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.478C14
;  CURRENT APPLICATION NUMBER: US/09/702,705
;  CURRENT FILING DATE:  2000-10-30
;  NUMBER OF SEQ ID NOS:  1833
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 320
;   LENGTH: 3132
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-702-705-320

Alignment Scores:
Pred. No.:              3.12          Length:          3132
Score:                  11.00         Matches:         11
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            0.77%         Indels:          0
DB:                     4             Gaps:            0

US-09-787-097-12 (1-1429) x US-09-702-705-320 (1-3132)

Qy          61 LeuLeuLeuLeuLeuLeuLeuProProPro 71
               |||||||||||||||||||||||||||||||
Db         521 CTGCTGCTGCTGCTGCTGTTGCTGCCGCCCCCG 489

RESULT 44
US-09-736-457-320/c
; Sequence 320, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Wang, Aijun
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.478C15
;  CURRENT APPLICATION NUMBER: US/09/736,457
;  CURRENT FILING DATE:  2000-12-13
;  NUMBER OF SEQ ID NOS:  1864
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 320
;   LENGTH: 3132
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-736-457-320

Alignment Scores:
Pred. No.:              3.12          Length:          3132
Score:                  11.00         Matches:         11
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            0.77%         Indels:          0
DB:                     4             Gaps:            0

US-09-787-097-12 (1-1429) x US-09-736-457-320 (1-3132)

Qy          61 LeuLeuLeuLeuLeuLeuLeuProProPro 71
               |||||||||||||||||||||||||||||||
Db         521 CTGCTGCTGCTGCTGCTGTTGCTGCCGCCCCCG 489

RESULT 45
US-09-614-124B-320/c
; Sequence 320, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
;  APPLICANT: Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.478C9
;  CURRENT APPLICATION NUMBER: US/09/614,124B
;  CURRENT FILING DATE:  2001-07-11
;  NUMBER OF SEQ ID NOS: 1668
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 320
;   LENGTH: 3132
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-614-124B-320

Alignment Scores:
Pred. No.:              3.12          Length:          3132
Score:                  11.00         Matches:         11
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            0.77%         Indels:          0
DB:                     4             Gaps:            0

US-09-787-097-12 (1-1429) x US-09-614-124B-320 (1-3132)
```

```
Qy          61 LeuLeuLeuLeuLeuLeuLeuProProPro 71
               |||||||||||||||||||||||||||||||
Db         521 CTGCTGCTGCTGCTGCTGTTGCTGCCGCCCCCG 489

RESULT 46
US-09-671-325-320/c
; Sequence 320, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Fan, Liqun
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.478C12
;  CURRENT APPLICATION NUMBER: US/09/671,325
;  CURRENT FILING DATE:  2000-09-26
;  NUMBER OF SEQ ID NOS:  1825
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 320
;   LENGTH: 3132
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-671-325-320

Alignment Scores:
Pred. No.:              3.12          Length:        3132
Score:                  11.00         Matches:       11
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            0.77%         Indels:        0
DB:                     4             Gaps:          0

US-09-787-097-12 (1-1429) x US-09-671-325-320 (1-3132)

Qy          61 LeuLeuLeuLeuLeuLeuLeuProProPro 71
               |||||||||||||||||||||||||||||||
Db         521 CTGCTGCTGCTGCTGCTGTTGCTGCCGCCCCCG 489

RESULT 47
US-09-589-184-320/c
; Sequence 320, Application US/09589184
; Patent No. 6686447
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.478C8
;  CURRENT APPLICATION NUMBER: US/09/589,184
;  CURRENT FILING DATE:  2000-06-05
;  NUMBER OF SEQ ID NOS:  827
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 320
;   LENGTH: 3132
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-589-184-320

Alignment Scores:
Pred. No.:              3.12          Length:        3132
Score:                  11.00         Matches:       11
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            0.77%         Indels:        0
DB:                     4             Gaps:          0

US-09-787-097-12 (1-1429) x US-09-589-184-320 (1-3132)

Qy          61 LeuLeuLeuLeuLeuLeuLeuProProPro 71
               |||||||||||||||||||||||||||||||
Db         521 CTGCTGCTGCTGCTGCTGTTGCTGCCGCCCCCG 489

RESULT 48
US-08-884-569A-1
; Sequence 1, Application US/08884569A
; Patent No. 6399326
; GENERAL INFORMATION:
;  APPLICANT: CHIANG, MING-KO
;  APPLICANT:  FLANAGAN, JOHN G.
;  TITLE OF INVENTION: RECEPTOR TYROSINE PHOSPHATASE, AND USES RELATED THERETO
;  FILE REFERENCE: HMV-020.01
;  CURRENT APPLICATION NUMBER: US/08/884,569A
;  CURRENT FILING DATE:  1997-06-27
;  PRIOR APPLICATION NUMBER: 60/021,040
;  PRIOR FILING DATE: 1996-07-02
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 3200
;   TYPE: DNA
;   ORGANISM: Mus sp.
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (33)..(3035)
US-08-884-569A-1

Alignment Scores:
Pred. No.:              3.18          Length:        3200
Score:                  11.00         Matches:       11
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            0.77%         Indels:        0
DB:                     4             Gaps:          0

US-09-787-097-12 (1-1429) x US-08-884-569A-1 (1-3200)

Qy          62 LeuLeuLeuLeuLeuLeuProProProLeu 72
               |||||||||||||||||||||||||||||||||
Db          51 CTTTGCTGCTGCTGCTGCTACCGCCGCCGCTG 83

RESULT 49
US-09-367-891A-5/c
; Sequence 5, Application US/09367891A
; Patent No. 6524816
; GENERAL INFORMATION:
;  APPLICANT: RASMUSSEN, PREBEN
;  TITLE OF INVENTION: EXPRESSION ELEMENT
;  FILE REFERENCE: 078883/0111
;  CURRENT APPLICATION NUMBER: US/09/367,891A
;  CURRENT FILING DATE:  1999-12-03
;  PRIOR APPLICATION NUMBER: PCT/IB98/00312
;  PRIOR FILING DATE: 1998-03-02
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 3311
;   TYPE: DNA
;   ORGANISM: Unknown Organism
;   FEATURE:
;   OTHER INFORMATION: Description of Unknown Organism: amyA gene
US-09-367-891A-5

Alignment Scores:
```

```
Pred. No.:              3.3             Length:         3311
Score:                  11.00           Matches:        11
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            0.77%           Indels:         0
DB:                     4               Gaps:           0

US-09-787-097-12 (1-1429) x US-09-367-891A-5 (1-3311)

Qy          61 LeuLeuLeuLeuLeuLeuLeuProProPro 71
               |||||||||||||||||||||||||||||||
Db        2963 CTGCTGCTGCTGCTGCTGCTGCCGCCGCCG 2931

RESULT 50
US-09-166-350-32
; Sequence 32, Application US/09166350A
; Patent No. 6440663
; GENERAL INFORMATION:
;  APPLICANT: Scanlan, Matthew
;  APPLICANT: Chen, Yao
;  APPLICANT: Stockert, Elisabeth
;  APPLICANT: Old, Lloyd
;  APPLICANT: Jager, Elke
;  APPLICANT: Knuth, Alex
;  TITLE OF INVENTION: Renal Cancer Associated Antigens and
;  TITLE OF INVENTION: Uses Therefor
;  FILE REFERENCE: L0461/7051
;  CURRENT APPLICATION NUMBER: US/09/166,350A
;  CURRENT FILING DATE: 1998-10-05
;  EARLIER APPLICATION NUMBER: US 09/166,350
;  EARLIER FILING DATE: 1998-10-05
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 32
;  LENGTH: 4169
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US-09-166-350-32

Alignment Scores:
Pred. No.:              4.15            Length:         4169
Score:                  11.00           Matches:        11
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            0.77%           Indels:         0
DB:                     4               Gaps:           0

US-09-787-097-12 (1-1429) x US-09-166-350-32 (1-4169)

Qy          61 LeuLeuLeuLeuLeuLeuLeuProProPro 71
               |||||||||||||||||||||||||||||||
Db          53 CTGCTGCTGCTGCTGCTACTGCTACCACCGCCT 85

RESULT 51
US-09-023-655-1386/c
; Sequence 1386, Application US/09023655
; Patent No. 6607879
;  GENERAL INFORMATION:
;    APPLICANT:  Cocks, Benjamin G.
;    APPLICANT:  Susan G. Stuart
;    APPLICANT:  Jeffrey J. Seilhamer
;    TITLE OF INVENTION:  COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
;    TITLE OF INVENTION:  EXPRESSION
;    NUMBER OF SEQUENCES:  1508
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/023,655
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Zeller, Karen J.
;      REGISTRATION NUMBER:  37,071
;      REFERENCE/DOCKET NUMBER:  PA-0001 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 855-0555
;      TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  1386:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5828 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GENBANK
;      CLONE:  g386256
US-09-023-655-1386

Alignment Scores:
Pred. No.:              5.8             Length:         5828
Score:                  11.00           Matches:        11
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            0.77%           Indels:         0
DB:                     4               Gaps:           0

US-09-787-097-12 (1-1429) x US-09-023-655-1386 (1-5828)

Qy          61 LeuLeuLeuLeuLeuLeuLeuProProPro 71
               |||||||||||||||||||||||||||||||
Db         392 CTGCTGCTGCTGCTGCTGCTGCCGCCGCCC 360

RESULT 52
US-10-027-983-11
; Sequence 11, Application US/10027983
; Patent No. 6617162
; GENERAL INFORMATION:
;  APPLICANT: Kenneth W. Dobie
;  APPLICANT:  Mark P. Roach
;  TITLE OF INVENTION: ANTISENSE MODULATION OF ESTROGEN RECEPTOR ALPHA EXPRESSION
;  FILE REFERENCE: RTS-0340
;  CURRENT APPLICATION NUMBER: US/10/027,983
;  CURRENT FILING DATE:  2001-12-18
;  NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 11
;  LENGTH: 392000
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY: unsure
;  LOCATION: 137740
;  OTHER INFORMATION: unknown
;  NAME/KEY: unsure
;  LOCATION: 137742
;  OTHER INFORMATION: unknown
;  NAME/KEY: misc_feature
;  LOCATION: (138122)...(138221)
;  OTHER INFORMATION: n = A,T,C or G
;  NAME/KEY: unsure
;  LOCATION: 145507
;  OTHER INFORMATION: unknown
;  NAME/KEY: unsure
```

```
;   LOCATION: 151967
;   OTHER INFORMATION: unknown
;   NAME/KEY: misc_feature
;   LOCATION: (151967)...(1542066)
;   OTHER INFORMATION: n = A,T,C or G
;   NAME/KEY: unsure
;   LOCATION: 154217
;   OTHER INFORMATION: unknown
;   NAME/KEY: misc_feature
;   LOCATION: (164037)...(164136)
;   OTHER INFORMATION: n = A,T,C or G
;   NAME/KEY: misc_feature
;   LOCATION: (174657)...(174756)
;   OTHER INFORMATION: n = A,T,C or G
;   NAME/KEY: misc_feature
;   LOCATION: (186224)...(186323)
;   OTHER INFORMATION: n = A,T,C or G
;   NAME/KEY: misc_feature
;   LOCATION: (195242)...(195341)
;   OTHER INFORMATION: n = A,T,C or G
;   NAME/KEY: unsure
;   LOCATION: 202703
;   OTHER INFORMATION: unknown
;   NAME/KEY: misc_feature
;   LOCATION: (202771)...(202870)
;   OTHER INFORMATION: n = A,T,C or G
;   NAME/KEY: misc_feature
;   LOCATION: (206246)...(215602)
;   OTHER INFORMATION: n = A,T,C or G
;   NAME/KEY: misc_feature
;   LOCATION: (218126)...(218225)
;   OTHER INFORMATION: n = A,T,C or G
;   NAME/KEY: misc_feature
;   LOCATION: (220360)...(220459)
;   OTHER INFORMATION: n = A,T,C or G
;   NAME/KEY: misc_feature
;   LOCATION: (222717)...(222816)
;   OTHER INFORMATION: n = A,T,C or G
;   NAME/KEY: misc_feature
;   LOCATION: (223981)...(224080)
;   OTHER INFORMATION: n = A,T,C or G
;   NAME/KEY: misc_feature
;   LOCATION: (227487)...(227586)
;   OTHER INFORMATION: n = A,T,C or G
;   NAME/KEY: misc_feature
;   LOCATION: (230157)...(230256)
;   OTHER INFORMATION: n = A,T,C or G
;   NAME/KEY: misc_feature
;   LOCATION: (232299)...(232398)
;   OTHER INFORMATION: n = A,T,C or G
;   NAME/KEY: misc_feature
;   LOCATION: (236552)...(2366651)
;   OTHER INFORMATION: n = A,T,C or G
;   NAME/KEY: misc_feature
;   LOCATION: (238789)...(248788)
;   OTHER INFORMATION: n = A,T,C or G
;   NAME/KEY: exon
;   LOCATION: (118288)...(119101)
;   OTHER INFORMATION: exon 1C
;   NAME/KEY: exon:intron junction
;   LOCATION: (151129)...(151130)
;   OTHER INFORMATION: exon 5:intron 5
;   NAME/KEY: exon:intron junction
;   LOCATION: (299248)...(299249)
;   OTHER INFORMATION: exon 9:intron 9
;   NAME/KEY: exon:intron junction
;   LOCATION: (348578)...(348579)
;   OTHER INFORMATION: exon 10:intron 10
;   NAME/KEY: intron
;   LOCATION: (348579)...(381838)
;   OTHER INFORMATION: intron 10
;   NAME/KEY: intron:exon junction
;   LOCATION: (386185)...(386186)
;   OTHER INFORMATION: intron 11:exon 12
US-10-027-983-11

Alignment Scores:
Pred. No.:                390              Length:          392000
Score:                    11.00            Matches:         11
Percent Similarity:       100.00%          Conservative:    0
Best Local Similarity:    100.00%          Mismatches:      0
Query Match:              0.77%            Indels:          0
DB:                       4                Gaps:            0

US-09-787-097-12 (1-1429) x US-10-027-983-11 (1-392000)

Qy          61 LeuLeuLeuLeuLeuLeuLeuProProPro 71
               ||||||||||||||||||||||||||||||
Db      298327 CTCCTCCTTCTCCTTCTTCTTCTTCCTCCTCCT 298359

RESULT 53
US-08-811-492-132/c
; Sequence 132, Application US/08811492
; Patent No. 5834247
;  GENERAL INFORMATION:
;    APPLICANT:  COMB, DONALD G.
;    APPLICANT:  PERLER, FRANCINE B.
;    APPLICANT:  JACK, WILLIAM E.
;    APPLICANT:  XU, MING-QUN
;    APPLICANT:  HODGES, ROBERT A.
;    APPLICANT:  NOREN, CHRISTOPHER J.
;    APPLICANT:  CHONG, SHAORONG S.C.
;    APPLICANT:  ADAM, ERIC
;    APPLICANT:  SOUTHWORTH, MAURICE
;    TITLE OF INVENTION:  MODIFIED PROTEINS, METHODS OF THEIR
;    TITLE OF INVENTION:  PRODUCTION AND METHODS FOR PURIFICATION OF TARGET
;    TITLE OF INVENTION:  PROTEINS
;    NUMBER OF SEQUENCES:  155
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  GREGORY D. WILLIAMS; NEW ENGLAND BIOLABS, INC.
;      STREET:  32 TOZER RAOD
;      CITY:  BEVERLY
;      STATE:  MASSACHUSETTS
;      COUNTRY:  USA
;      ZIP:  01915
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC\_DOS/MS\_DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/811,492
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/580,555
;      FILING DATE:  29-DEC-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/496,247
;      FILING DATE:  28-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/146,885
;      FILING DATE:  03-NOV-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/004,139
;      FILING DATE:  09-DEC-1992
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams, Gregory D
;      REGISTRATION NUMBER:  30901
;      REFERENCE/DOCKET NUMBER:  NEB-036C4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  508-927-5054
```

```
;      TELEFAX:  509-927-1705
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  132:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  100 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
US-08-811-492-132

Alignment Scores:
Pred. No.:                0.859          Length:         100
Score:                    10.00          Matches:        10
Percent Similarity:       100.00%        Conservative:   0
Best Local Similarity:    100.00%        Mismatches:     0
Query Match:              0.70%          Indels:         0
DB:                       2              Gaps:           0

US-09-787-097-12 (1-1429) x US-08-811-492-132 (1-100)

Qy         59 ProArgLeuLeuLeuLeuLeuLeuLeu 68
              |||||||||||||||||||||||||||||
Db         42 CCGAGGTTGTTGTTATTGTTATTGTTGTTG 13

RESULT 54
US-08-811-492-133/c
; Sequence 133, Application US/08811492
; Patent No. 5834247
;  GENERAL INFORMATION:
;    APPLICANT:  COMB, DONALD G.
;    APPLICANT:  PERLER, FRANCINE B.
;    APPLICANT:  JACK, WILLIAM E.
;    APPLICANT:  XU, MING-QUN
;    APPLICANT:  HODGES, ROBERT A.
;    APPLICANT:  NOREN, CHRISTOPHER J.
;    APPLICANT:  CHONG, SHAORONG S.C.
;    APPLICANT:  ADAM, ERIC
;    APPLICANT:  SOUTHWORTH, MAURICE
;    TITLE OF INVENTION:  MODIFIED PROTEINS, METHODS OF THEIR
;    TITLE OF INVENTION:  PRODUCTION AND METHODS FOR PURIFICATION OF TARGET
;    TITLE OF INVENTION:  PROTEINS
;    NUMBER OF SEQUENCES:  155
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  GREGORY D. WILLIAMS; NEW ENGLAND BIOLABS, INC.
;      STREET:  32 TOZER RAOD
;      CITY:  BEVERLY
;      STATE:  MASSACHUSETTS
;      COUNTRY:  USA
;      ZIP:  01915
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC\_DOS/MS\_DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/811,492
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/580,555
;      FILING DATE:  29-DEC-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/496,247
;      FILING DATE:  28-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/146,885
;      FILING DATE:  03-NOV-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/004,139
;      FILING DATE:  09-DEC-1992
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams, Gregory D
;      REGISTRATION NUMBER:  30901
;      REFERENCE/DOCKET NUMBER:  NEB-036C4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  508-927-5054
;      TELEFAX:  509-927-1705
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  133:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  109 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
US-08-811-492-133

Alignment Scores:
Pred. No.:                0.936          Length:         109
Score:                    10.00          Matches:        10
Percent Similarity:       100.00%        Conservative:   0
Best Local Similarity:    100.00%        Mismatches:     0
Query Match:              0.70%          Indels:         0
DB:                       2              Gaps:           0

US-09-787-097-12 (1-1429) x US-08-811-492-133 (1-109)

Qy         59 ProArgLeuLeuLeuLeuLeuLeuLeu 68
              |||||||||||||||||||||||||||||
Db         81 CCGAGGTTGTTGTTATTGTTATTGTTGTTG 52

RESULT 55
US-08-478-312-55/c
; Sequence 55, Application US/08478312
; Patent No. 5654276
;  GENERAL INFORMATION:
;    APPLICANT:  Barrett, Ronald W.
;    APPLICANT:  England, Bruce
;    APPLICANT:  Schatz, Peter
;    APPLICANT:  Sloan, Derek
;    APPLICANT:  Chen, Min-Jia
;    TITLE OF INVENTION:  Peptides and Compounds That Bind to the IL-5
;    TITLE OF INVENTION:  Receptor
;    NUMBER OF SEQUENCES:  59
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Affymax Technologies, N.V.
;      STREET:  4001 Miranda Ave.
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/478,312
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stevens, Lauren L.
;      REGISTRATION NUMBER:  36,691
;      REFERENCE/DOCKET NUMBER:  1088.1A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-496-2300
;      TELEFAX:  415-424-0832
;  INFORMATION FOR SEQ ID NO:  55:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  122 base pairs
;      TYPE:  nucleic acid
```

```
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (oligonucleotide)
US-08-478-312-55

Alignment Scores:
Pred. No.:                1.05          Length:          122
Score:                    10.00         Matches:         10
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              0.70%         Indels:          0
DB:                       1             Gaps:            0

US-09-787-097-12 (1-1429) x US-08-478-312-55 (1-122)

Qy          59 ProArgLeuLeuLeuLeuLeuLeuLeu 68
               ||||||||||||||||||||||||||||||
Db          53 CCGAGGTTGTTGTTATTGTTATTGTTGTTG 24

RESULT 56
US-08-485-302-55/c
; Sequence 55, Application US/08485302
; Patent No. 5668110
;  GENERAL INFORMATION:
;    APPLICANT:  Barrett, Ronald W.
;    APPLICANT:  England, Bruce
;    APPLICANT:  Schatz, Peter
;    APPLICANT:  Sloan, Derek
;    APPLICANT:  Chen, Min-Jia
;    TITLE OF INVENTION:  Peptides and Compounds That Bind to the IL-5
;    TITLE OF INVENTION:  Receptor
;    NUMBER OF SEQUENCES:  59
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Affymax Technologies, N.V.
;      STREET:  4001 Miranda Ave.
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/485,302
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stevens, Lauren L.
;      REGISTRATION NUMBER:  36,691
;      REFERENCE/DOCKET NUMBER:  1088.1B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-496-2300
;      TELEFAX:  415-424-0832
;  INFORMATION FOR SEQ ID NO:  55:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  122 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (oligonucleotide)
US-08-485-302-55

Alignment Scores:
Pred. No.:                1.05          Length:          122
Score:                    10.00         Matches:         10
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              0.70%         Indels:          0
DB:                       1             Gaps:            0

US-09-787-097-12 (1-1429) x US-08-485-302-55 (1-122)

Qy          59 ProArgLeuLeuLeuLeuLeuLeuLeu 68
               ||||||||||||||||||||||||||||||
Db          53 CCGAGGTTGTTGTTATTGTTATTGTTGTTG 24

RESULT 57
US-08-476-169-60/c
; Sequence 60, Application US/08476169
; Patent No. 5677280
;  GENERAL INFORMATION:
;    APPLICANT:  Barrett, Ronald W.
;    APPLICANT:  England, Bruce
;    APPLICANT:  Schatz, Peter
;    APPLICANT:  Sloan, Derek
;    APPLICANT:  Chen, Min-Jia
;    TITLE OF INVENTION:  Peptides and Compounds That Bind to the IL-5
;    TITLE OF INVENTION:  Receptor
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Affymax Technologies, N.V.
;      STREET:  4001 Miranda Ave.
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/476,169
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stevens, Lauren L.
;      REGISTRATION NUMBER:  36,691
;      REFERENCE/DOCKET NUMBER:  1088.2A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-496-2300
;      TELEFAX:  415-424-0832
;  INFORMATION FOR SEQ ID NO:  60:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  122 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (oligonucleotide)
US-08-476-169-60

Alignment Scores:
Pred. No.:                1.05          Length:          122
Score:                    10.00         Matches:         10
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              0.70%         Indels:          0
DB:                       1             Gaps:            0

US-09-787-097-12 (1-1429) x US-08-476-169-60 (1-122)

Qy          59 ProArgLeuLeuLeuLeuLeuLeuLeu 68
               ||||||||||||||||||||||||||||||
Db          53 CCGAGGTTGTTGTTATTGTTATTGTTGTTG 24

RESULT 58
US-08-484-083-60/c
; Sequence 60, Application US/08484083
; Patent No. 5683983
;  GENERAL INFORMATION:
;    APPLICANT:  Barrett, Ronald W.
;    APPLICANT:  England, Bruce
;    APPLICANT:  Schatz, Peter
;    APPLICANT:  Sloan, Derek
```

```
;    APPLICANT:  Chen, Min-Jia
;    TITLE OF INVENTION:  Peptides and Compounds That Bind to the IL-5
;    TITLE OF INVENTION:  Receptor
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Affymax Technologies, N.V.
;      STREET:  4001 Miranda Ave.
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,083
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stevens, Lauren L.
;      REGISTRATION NUMBER:  36,691
;      REFERENCE/DOCKET NUMBER:  1088.2B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-496-2300
;      TELEFAX:  415-424-0832
;  INFORMATION FOR SEQ ID NO:  60:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  122 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (oligonucleotide)
US-08-484-083-60

Alignment Scores:
Pred. No.:                1.05          Length:          122
Score:                    10.00         Matches:         10
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              0.70%         Indels:          0
DB:                       1             Gaps:            0

US-09-787-097-12 (1-1429) x US-08-484-083-60 (1-122)

Qy         59 ProArgLeuLeuLeuLeuLeuLeuLeu 68
              |||||||||||||||||||||||||||||
Db         53 CCGAGGTTGTTGTTATTGTTATTGTTGTTG 24

RESULT 59
US-08-764-640-237/c
; Sequence 237, Application US/08764640
; Patent No. 5869451
; Patent No. 5869451 5837683
;  GENERAL INFORMATION:
;    APPLICANT:  Dower, William J.
;    APPLICANT:  Barrett, Ronald W.
;    APPLICANT:  Cwirla, Steven E.
;    APPLICANT:  Gates, Christian
;    APPLICANT:  Schatz, Peter J.
;    APPLICANT:  Balasubramanian, Palaniappan
;    APPLICANT:  Wagstrom, Christopher R.
;    APPLICANT:  Hendren, Richard W.
;    APPLICANT:  Deprince, Randolph B.
;    APPLICANT:  Podduturi, Surekha
;    APPLICANT:  Yin, Qun
;    TITLE OF INVENTION:  PEPTIDES AND COMPOUNDS THAT BIND TO A
;    TITLE OF INVENTION:  RECEPTOR
;    NUMBER OF SEQUENCES:  244
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Glaxo Wellcome
;      STREET:  Five Moore Drive, P.O. Box 13398
;      CITY:  Research Triangle Park
;      STATE:  NC
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/764,640
;      FILING DATE:  11-DEC-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hrubiec, Robert T.
;      REGISTRATION NUMBER:  36,392
;      REFERENCE/DOCKET NUMBER:  PK3281
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-248-1000
;  INFORMATION FOR SEQ ID NO:  237:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  122 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..120
US-08-764-640-237

Alignment Scores:
Pred. No.:                1.05          Length:          122
Score:                    10.00         Matches:         10
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              0.70%         Indels:          0
DB:                       2             Gaps:            0

US-09-787-097-12 (1-1429) x US-08-764-640-237 (1-122)

Qy         59 ProArgLeuLeuLeuLeuLeuLeuLeu 68
              |||||||||||||||||||||||||||||
Db         53 CCGAGGTTGTTGTTATTGTTATTGTTGTTG 24

RESULT 60
US-08-973-225-225/c
; Sequence 225, Application US/08973225A
; Patent No. 6083913
;   GENERAL INFORMATION:
;        APPLICANT: Dower, William J.
;                   Barrett, Ronald W.
;                   Cwirla, Steven E.
;                   Duffin, David J.
;                   Gates, Christian
;                   Haselden, Sherril S.
;                   Mattheakis, Larry C.
;                   Schatz, Peter J.
;                   Wagstrom, Christopher R.
;                   Wrighton, Nicholas C.
;        TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A
;                            THROMBOPOIETIN RECEPTOR
;        NUMBER OF SEQUENCES: 232
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Glaxo Wellcome
;             STREET: Five Moore Drive, P.O. Box 13398
;             CITY: Research Triangle Park
;             STATE: NC
;             COUNTRY: USA
;             ZIP: 27709
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
```

```
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;          CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/973,225A
;             FILING DATE: 04-Dec-1997
;          ATTORNEY/AGENT INFORMATION:
;             NAME: Hrubiec, Robert T.
;             REGISTRATION NUMBER: 36,392
;             REFERENCE/DOCKET NUMBER: PK3065USW
;          TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 919-248-1000
;    INFORMATION FOR SEQ ID NO: 225:
;          SEQUENCE CHARACTERISTICS:
;             LENGTH: 122 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;          MOLECULE TYPE: cDNA
;          FEATURE:
;             NAME/KEY:  CDS
;             LOCATION:  1..120
;          SEQUENCE DESCRIPTION: SEQ ID NO: 225:
US-08-973-225-225

Alignment Scores:
Pred. No.:                1.05          Length:          122
Score:                    10.00         Matches:         10
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              0.70%         Indels:          0
DB:                       3             Gaps:            0

US-09-787-097-12 (1-1429) x US-08-973-225-225 (1-122)

Qy          59 ProArgLeuLeuLeuLeuLeuLeuLeuLeu 68
               ||||||||||||||||||||||||||||||||
Db          53 CCGAGGTTGTTGTTATTGTTATTGTTGTTG 24

RESULT 61
US-09-244-298A-237/c
; Sequence 237, Application US/09244298A
; Patent No. 6121238
;  GENERAL INFORMATION:
;    APPLICANT:  Dower, William J.
;    APPLICANT:  Barrett, Ronald W.
;    APPLICANT:  Cwirla, Steven E.
;    APPLICANT:  Gates, Christian
;    APPLICANT:  Schatz, Peter J.
;    APPLICANT:  Balasubramanian, Palaniappan
;    APPLICANT:  Wagstrom, Christopher R.
;    APPLICANT:  Hendren, Richard W.
;    APPLICANT:  Deprince, Randolph B.
;    APPLICANT:  Podduturi, Surekha
;    APPLICANT:  Yin, Qun
;    TITLE OF INVENTION:  PEPTIDES AND COMPOUNDS THAT BIND TO A
;    TITLE OF INVENTION:  RECEPTOR
;    NUMBER OF SEQUENCES:  244
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Glaxo Wellcome
;      STREET:  Five Moore Drive, P.O. Box 13398
;      CITY:  Research Triangle Park
;      STATE:  NC
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/244,298A
;      FILING DATE:  11-DEC-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hrubiec, Robert T.
;      REGISTRATION NUMBER:  36,392
;      REFERENCE/DOCKET NUMBER:  PK3281
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-248-1000
;  INFORMATION FOR SEQ ID NO:  237:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  122 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..120
US-09-244-298A-237

Alignment Scores:
Pred. No.:                1.05          Length:          122
Score:                    10.00         Matches:         10
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              0.70%         Indels:          0
DB:                       3             Gaps:            0

US-09-787-097-12 (1-1429) x US-09-244-298A-237 (1-122)

Qy          59 ProArgLeuLeuLeuLeuLeuLeuLeuLeu 68
               ||||||||||||||||||||||||||||||||
Db          53 CCGAGGTTGTTGTTATTGTTATTGTTGTTG 24

RESULT 62
US-09-516-704-237/c
; Sequence 237, Application US/09516704
; Patent No. 6251864
;  GENERAL INFORMATION:
;       APPLICANT: Dower, William J.
;                  Barrett, Ronald W.
;                  Cwirla, Steven E.
;                  Gates, Christian
;                  Schatz, Peter J.
;                  Balasubramanian, Palaniappan
;                  Wagstrom, Christopher R.
;                  Hendren, Richard W.
;                  Deprince, Randolph B.
;                  Podduturi, Surekha
;       TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A
;                           RECEPTOR
;       NUMBER OF SEQUENCES: 244
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Glaxo Wellcome
;            STREET: Five Moore Drive, P.O. Box 13398
;            CITY: Research Triangle Park
;            STATE: NC
;            COUNTRY: USA
;            ZIP: 27709
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/516,704
;            FILING DATE: 01-Mar-2000
;            CLASSIFICATION: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Hrubiec, Robert T.
;            REGISTRATION NUMBER: 36,392
;            REFERENCE/DOCKET NUMBER: PK3281
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 919-248-1000
;  INFORMATION FOR SEQ ID NO: 237:
```

```
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 122 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        FEATURE:
;             NAME/KEY:  CDS
;             LOCATION:  1..120
;        SEQUENCE DESCRIPTION: SEQ ID NO: 237:
US-09-516-704-237

Alignment Scores:
Pred. No.:              1.05          Length:        122
Score:                  10.00         Matches:       10
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            0.70%         Indels:        0
DB:                     3             Gaps:          0

US-09-787-097-12 (1-1429) x US-09-516-704-237 (1-122)

Qy         59 ProArgLeuLeuLeuLeuLeuLeuLeu 68
              |||||||||||||||||||||||||||||
Db         53 CCGAGGTTGTTGTTATTGTTATTGTTGTTG 24

RESULT 63
US-09-549-090-225/c
; Sequence 225, Application US/09549090
; Patent No. 6465430
;   GENERAL INFORMATION:
;        APPLICANT: Dower, William J.
;                   Barrett, Ronald W.
;                   Cwirla, Steven E.
;                   Duffin, David J.
;                   Gates, Christian
;                   Haselden, Sherril S.
;                   Mattheakis, Larry C.
;                   Schatz, Peter J.
;                   Wagstrom, Christopher R.
;                   Wrighton, Nicholas C.
;        TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A
;                            THROMBOPOIETIN RECEPTOR
;        NUMBER OF SEQUENCES: 232
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Glaxo Wellcome
;             STREET: Five Moore Drive, P.O. Box 13398
;             CITY: Research Triangle Park
;             STATE: NC
;             COUNTRY: USA
;             ZIP: 27709
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/549,090
;             FILING DATE: 13-Apr-2000
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US 08/973,225
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Hrubiec, Robert T.
;             REGISTRATION NUMBER: 36,392
;             REFERENCE/DOCKET NUMBER: PK3065USW
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 919-248-1000
;   INFORMATION FOR SEQ ID NO: 225:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 122 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        FEATURE:
;             NAME/KEY:  CDS
;             LOCATION:  1..120
;        SEQUENCE DESCRIPTION: SEQ ID NO: 225:
US-09-549-090-225

Alignment Scores:
Pred. No.:              1.05          Length:        122
Score:                  10.00         Matches:       10
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            0.70%         Indels:        0
DB:                     4             Gaps:          0

US-09-787-097-12 (1-1429) x US-09-549-090-225 (1-122)

Qy         59 ProArgLeuLeuLeuLeuLeuLeuLeu 68
              |||||||||||||||||||||||||||||
Db         53 CCGAGGTTGTTGTTATTGTTATTGTTGTTG 24

RESULT 64
US-09-832-230A-237/c
; Sequence 237, Application US/09832230A
; Patent No. 6506362
;   GENERAL INFORMATION:
;        APPLICANT: Dower, William J. et al
;        TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A
;                            RECEPTOR
;        NUMBER OF SEQUENCES: 244
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Glaxo Wellcome
;             STREET: Five Moore Drive, P.O. Box 13398
;             CITY: Research Triangle Park
;             STATE: NC
;             COUNTRY: USA
;             ZIP: 27709
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/832,230A
;             FILING DATE: 10-Apr-2001
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Hrubiec, Robert T.
;             REGISTRATION NUMBER: 36,392
;             REFERENCE/DOCKET NUMBER: PK3281
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 919-248-1000
;   INFORMATION FOR SEQ ID NO: 237:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 122 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        FEATURE:
;             NAME/KEY:  CDS
;             LOCATION:  1..120
;        SEQUENCE DESCRIPTION: SEQ ID NO: 237:
US-09-832-230A-237

Alignment Scores:
Pred. No.:              1.05          Length:        122
Score:                  10.00         Matches:       10
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            0.70%         Indels:        0
DB:                     4             Gaps:          0
```

US-09-787-097-12 (1-1429) x US-09-832-230A-237 (1-122)

```
Qy         59 ProArgLeuLeuLeuLeuLeuLeuLeu 68
              |||||||||||||||||||||||||||||
Db         53 CCGAGGTTGTTGTTATTGTTATTGTTGTTG 24
```

RESULT 65
US-08-938-548B-21
; Sequence 21, Application US/08938548B
; Patent No. 6001963
; GENERAL INFORMATION:
; APPLICANT: Yanagisawa, Masashi
; APPLICANT: Bergsma, Derk
; APPLICANT: Wilson, Shelagh
; APPLICANT: Brooks, David
; APPLICANT: Gellai, Miklos
; TITLE OF INVENTION: NOVEL LIGANDS OF THE NEUROPEPTIDE
; TITLE OF INVENTION: RECEPTOR HFGAN72
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: United States of America
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/938,548B
; FILING DATE: 26-SEPT-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/887,382
; FILING DATE: 2-JUL-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/820,519
; FILING DATE: 19-MAR-1997
; APPLICATION NUMBER: 60/033,604
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Elizabeth J. Hecht
; REGISTRATION NUMBER: 41,824
; REFERENCE/DOCKET NUMBER: ATG50037-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5009
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 577 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-938-548B-21

Alignment Scores:

Pred. No.:	4.95	Length:	577
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.70%	Indels:	0
DB:	3	Gaps:	0

US-09-787-097-12 (1-1429) x US-08-938-548B-21 (1-577)

```
Qy         61 LeuLeuLeuLeuLeuLeuLeuProPro 70
              ||||||||||||||||||||||||||||||
```

```
Db        130 CTACTGCTGCTGCTGCTGCTGCCGCCC 159
```

RESULT 66
US-08-939-093A-21
; Sequence 21, Application US/08939093A
; Patent No. 6309854
; GENERAL INFORMATION:
; APPLICANT: Yanagisawa, Masashi
; APPLICANT: Bergsma, Derk
; APPLICANT: Wilson, Shelagh
; APPLICANT: Brooks, David
; APPLICANT: Gellai, Miklos
; TITLE OF INVENTION: NOVEL LIGANDS OF THE NEUROPEPTIDE
; TITLE OF INVENTION: RECEPTOR HFGAN72
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: United States of America
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/939,093A
; FILING DATE: 26-SEPT-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/887,382
; FILING DATE: 2-JUL-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/820,519
; FILING DATE: 19-MAR-1997
; APPLICATION NUMBER: 60/033,604
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: King, William T.
; REGISTRATION NUMBER: 30,954
; REFERENCE/DOCKET NUMBER: ATG50037-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5219
; TELEFAX: 610-270-4026
; TELEX:
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 577 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-939-093A-21

Alignment Scores:

Pred. No.:	4.95	Length:	577
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.70%	Indels:	0
DB:	4	Gaps:	0

US-09-787-097-12 (1-1429) x US-08-939-093A-21 (1-577)

```
Qy         61 LeuLeuLeuLeuLeuLeuLeuProPro 70
              ||||||||||||||||||||||||||||||
Db        130 CTACTGCTGCTGCTGCTGCTGCCGCCC 159
```

RESULT 67
US-09-211-823C-21
; Sequence 21, Application US/09211823C

```
; Patent No. 6664229
; GENERAL INFORMATION:
;  APPLICANT:  HAGEN, JAMES JOSEPH
;  APPLICANT:  TERRETT, JONATHAN ALEXANDER
;  APPLICANT:  UPTON, NEIL
;  APPLICANT:  PIPER, DAVID
;  APPLICANT:  SMITH, MARTIN IAN
;  APPLICANT:  KENNETT, GUY ANTHONY
;  APPLICANT:  PATEL, SARASWATI R.
;  TITLE OF INVENTION:  METHODS OF TREATMENT USING NOVEL LIGANDS
;  TITLE OF INVENTION:  OF THE NEUROPEPTIDE RECEPTOR HFGAN72 AND AGONISTS OR
;  TITLE OF INVENTION:  ANTAGONISTS THEREOF
;  FILE REFERENCE:  P50745
;  CURRENT APPLICATION NUMBER: US/09/211,823C
;  CURRENT FILING DATE:  1998-12-15
;  PRIOR APPLICATION NUMBER:  US 60/069,459
;  PRIOR FILING DATE:  1997-12-15
;  PRIOR APPLICATION NUMBER:  US 60/069,785
;  PRIOR FILING DATE:  1997-12-16
;  NUMBER OF SEQ ID NOS:  23
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 21
;   LENGTH: 577
;   TYPE: DNA
;   ORGANISM: HOMO SAPIENS
US-09-211-823C-21

Alignment Scores:
Pred. No.:                4.95          Length:          577
Score:                    10.00         Matches:         10
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              0.70%         Indels:          0
DB:                       4             Gaps:            0

US-09-787-097-12 (1-1429) x US-09-211-823C-21 (1-577)

Qy          61 LeuLeuLeuLeuLeuLeuLeuProPro 70
               ||||||||||||||||||||||||||||||
Db         130 CTACTGCTGCTGCTGCTGCTGCCGCCC 159

RESULT 68
US-09-325-932A-31/c
; Sequence 31, Application US/09325932A
; Patent No. 6451604
; GENERAL INFORMATION:
;  APPLICANT:  Flinn, Barry
;  APPLICANT:  Lasham, Annette
;  TITLE OF INVENTION:  Compositions affecting programmed cell
;  TITLE OF INVENTION:  death and their use in the modification of forestry plant develc
;  FILE REFERENCE:  1022
;  CURRENT APPLICATION NUMBER: US/09/325,932A
;  CURRENT FILING DATE:  1999-06-04
;  NUMBER OF SEQ ID NOS:  206
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 31
;   LENGTH: 614
;   TYPE: DNA
;   ORGANISM: Pinus radiata
US-09-325-932A-31

Alignment Scores:
Pred. No.:                5.27          Length:          614
Score:                    10.00         Matches:         10
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              0.70%         Indels:          0
DB:                       4             Gaps:            0

US-09-787-097-12 (1-1429) x US-09-325-932A-31 (1-614)

Qy          61 LeuLeuLeuLeuLeuLeuLeuProPro 70
               ||||||||||||||||||||||||||||||
Db         329 CTCCTCCTCCTCCTCCTCTTGCCCCCC 300

RESULT 69
US-09-833-381-782
; Sequence 782, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
;  APPLICANT: Robison, Keith E.
;  TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
;  FILE REFERENCE: 5800-119
;  CURRENT APPLICATION NUMBER: US/09/833,381
;  CURRENT FILING DATE:  2001-04-11
;  PRIOR APPLICATION NUMBER: 09/516,448
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 2050
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 782
;   LENGTH: 867
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(867)
;   OTHER INFORMATION: n = A,T,C or G
US-09-833-381-782

Alignment Scores:
Pred. No.:                7.44          Length:          867
Score:                    10.00         Matches:         10
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              0.70%         Indels:          0
DB:                       4             Gaps:            0

US-09-787-097-12 (1-1429) x US-09-833-381-782 (1-867)

Qy          44 GlyAlaGlyLeuArgLeuProArgLeuLeu 53
               ||||||||||||||||||||||||||||||
Db          50 GGTGCGGGCCTGCGCCTCCCTCGGCTCCTG 79

RESULT 70
US-08-811-481-26
; Sequence 26, Application US/08811481
; Patent No. 6300093
;  GENERAL INFORMATION:
;    APPLICANT:  Kindsvogel, Wayne
;    APPLICANT:  Jelinek, Laura J.
;    APPLICANT:  Sheppard, Paul O.
;    APPLICANT:  Hagopian, William A.
;    APPLICANT:  LaGasse, James M.
;    TITLE OF INVENTION:  ISLET CELL ANTIGEN 1851
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ZymoGenetics, Inc.
;      STREET:  1201 Eastlake Avenue East
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98102
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/811,481
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lingenfelter, Susan
```

; REGISTRATION NUMBER: P-41,156
; REFERENCE/DOCKET NUMBER: 95-36
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6675
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 999 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-811-481-26

Alignment Scores:

Pred. No.:	8.57	Length:	999
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.70%	Indels:	0
DB:	4	Gaps:	0

US-09-787-097-12 (1-1429) x US-08-811-481-26 (1-999)

```
Qy         61 LeuLeuLeuLeuLeuLeuLeuProPro 70
              ||||||||||||||||||||||||||||
Db          1 CTGTTGCTGCTACTGCTGCTGCTGCCGCCA 30
```

RESULT 71
US-09-876-527-26
; Sequence 26, Application US/09876527
; Patent No. 6627735
; GENERAL INFORMATION:
; APPLICANT: Kindsvogel, Wayne
; Jelinek, Laura J.
; Sheppard, Paul O.
; Hagopian, William A.
; LaGasse, James M.
; TITLE OF INVENTION: ISLET CELL ANTIGEN 1851
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/876,527
; FILING DATE: 07-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/811,481
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Lingenfelter, Susan
; REGISTRATION NUMBER: P-41,156
; REFERENCE/DOCKET NUMBER: 95-36
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6675
; TELEFAX: 206-442-6678
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 999 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-876-527-26

Alignment Scores:

Pred. No.:	8.57	Length:	999
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.70%	Indels:	0
DB:	4	Gaps:	0

US-09-787-097-12 (1-1429) x US-09-876-527-26 (1-999)

```
Qy         61 LeuLeuLeuLeuLeuLeuLeuProPro 70
              ||||||||||||||||||||||||||||
Db          1 CTGTTGCTGCTACTGCTGCTGCTGCCGCCA 30
```

RESULT 72
US-08-811-481-27
; Sequence 27, Application US/08811481
; Patent No. 6300093
; GENERAL INFORMATION:
; APPLICANT: Kindsvogel, Wayne
; APPLICANT: Jelinek, Laura J.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Hagopian, William A.
; APPLICANT: LaGasse, James M.
; TITLE OF INVENTION: ISLET CELL ANTIGEN 1851
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,481
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lingenfelter, Susan
; REGISTRATION NUMBER: P-41,156
; REFERENCE/DOCKET NUMBER: 95-36
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6675
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1011 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-811-481-27

Alignment Scores:

Pred. No.:	8.68	Length:	1011
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.70%	Indels:	0

```
DB:                    4              Gaps:          0

US-09-787-097-12 (1-1429) x US-08-811-481-27 (1-1011)

Qy         61 LeuLeuLeuLeuLeuLeuLeuProPro 70
              ||||||||||||||||||||||||||||||
Db         13 CTGTTGCTGCTACTGCTGCTGCTGCCGCCA 42

RESULT 73
US-09-876-527-27
; Sequence 27, Application US/09876527
; Patent No. 6627735
;   GENERAL INFORMATION:
;        APPLICANT: Kindsvogel, Wayne
;                   Jelinek, Laura J.
;                   Sheppard, Paul O.
;                   Hagopian, William A.
;                   LaGasse, James M.
;        TITLE OF INVENTION: ISLET CELL ANTIGEN 1851
;        NUMBER OF SEQUENCES: 34
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: ZymoGenetics, Inc.
;             STREET: 1201 Eastlake Avenue East
;             CITY: Seattle
;             STATE: WA
;             COUNTRY: USA
;             ZIP: 98102
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: DOS
;             SOFTWARE: FastSEQ for Windows Version 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/876,527
;             FILING DATE: 07-Jun-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/811,481
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Lingenfelter, Susan
;             REGISTRATION NUMBER: P-41,156
;             REFERENCE/DOCKET NUMBER: 95-36
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 206-442-6675
;             TELEFAX: 206-442-6678
;             TELEX: <Unknown>
;   INFORMATION FOR SEQ ID NO: 27:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1011 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-09-876-527-27

Alignment Scores:
Pred. No.:              8.68          Length:        1011
Score:                  10.00         Matches:       10
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            0.70%         Indels:        0
DB:                     4             Gaps:          0

US-09-787-097-12 (1-1429) x US-09-876-527-27 (1-1011)

Qy         61 LeuLeuLeuLeuLeuLeuLeuProPro 70
              ||||||||||||||||||||||||||||||
Db         13 CTGTTGCTGCTACTGCTGCTGCTGCCGCCA 42

RESULT 74
US-09-489-039A-3636/c
; Sequence 3636, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
;  APPLICANT:  Gary Breton et. al
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
;  TITLE OF INVENTION:  PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  2709.2004001
;  CURRENT APPLICATION NUMBER: US/09/489,039A
;  CURRENT FILING DATE:  2000-01-27
;  PRIOR APPLICATION NUMBER:  US 60/117,747
;  PRIOR FILING DATE:  1999-01-29
;  NUMBER OF SEQ ID NOS:  14342
; SEQ ID NO 3636
;   LENGTH: 1146
;   TYPE: DNA
;   ORGANISM: Klebsiella pneumoniae
US-09-489-039A-3636

Alignment Scores:
Pred. No.:              9.84          Length:        1146
Score:                  10.00         Matches:       10
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            0.70%         Indels:        0
DB:                     4             Gaps:          0

US-09-787-097-12 (1-1429) x US-09-489-039A-3636 (1-1146)

Qy         89 AlaValSerGlySerAlaAlaAlaGluAla 98
              ||||||||||||||||||||||||||||||
Db        521 GCCGTCAGCGGCAGCGCGGCGGCGGAGGCA 492

RESULT 75
US-09-177-325-2/c
; Sequence 2, Application US/09177325B
; Patent No. 6214983
; GENERAL INFORMATION:
;  APPLICANT:      Tureci, Ozlem
;  APPLICANT:  Sakin, Ugur
;  APPLICANT:  Pfreundschuh, Michael
;  TITLE OF INVENTION:      Isolated Nucleic Acid Molecules Which Encode SCP Proteins, A
;  TITLE OF INVENTION:  Thereof
;  FILE REFERENCE:      LUD 5525
;  CURRENT APPLICATION NUMBER: US/09/177,325B
;  CURRENT FILING DATE:  1998-10-22
;  NUMBER OF SEQ ID NOS:      15
; SEQ ID NO 2
;   LENGTH: 1276
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
US-09-177-325-2

Alignment Scores:
Pred. No.:              11            Length:        1276
Score:                  10.00         Matches:       10
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            0.70%         Indels:        0
DB:                     3             Gaps:          0

US-09-787-097-12 (1-1429) x US-09-177-325-2 (1-1276)

Qy         61 LeuLeuLeuLeuLeuLeuLeuProPro 70
              ||||||||||||||||||||||||||||||
Db        813 CTCCTTCTTCTTCTCCTTCTTCTTCCTCCT 784


Search completed: March  1, 2004, 13:06:12
Job time : 751 secs
```